# Longitudinal Studies: Context, Measures, Construction and tools

John Milam, Ph. D.

This document was prepared with support from an Illinois Board of Higher Education (IBHE) grant provided to DePaul University for the purpose of providing technical assistance to institutions participating in the Illinois Higher Education Consortium (IHEC) and the Illinois Longitudinal Data System (ILDS). The views expressed represent the professional and personal views of the author and do not necessarily represent the views of IBHE or DePaul University.

## ACKNOWLEDGEMENTS

This document was prepared with support from an Illinois Board of Higher Education (IBHE) grant provided to DePaul University for the purpose of providing technical assistance to institutions participating in the Illinois Higher Education Consortium (IHEC) and the Illinois Longitudinal Data System (ILDS). The views expressed represent the professional and personal views of the author and do not necessarily represent the views of IBHE or DePaul University.

The author would like to thank Gerry McLaughlin for his encouragement of this monograph and kind edits, Brian Prescott for helpful comments on an early draft, Susan Doyle for grant support, and Lorraine Halsted for desktop publishing assistance. Thank you also to Lord Fairfax Community College and the Association for Institutional Research for support of this work and this digital edition.

# Table of Contents

# INTRODUCTION

## ORGANIZATION OF THE REPORT

The state of Illinois is in the process of building a state level student unit record data system that will ultimately track individual students from pre-school to their employment. This P-20 (Pre-Kindergarten to Graduate School) data system is called the Illinois Longitudinal Data System (ILDS). The purpose of this system is to facilitate achieving the state goals identified in the Illinois Public Agenda:

1. Increase educational attainment to match best-performing U.S. states and world countries;
2. Ensure college affordability for students, families, and taxpayers;
3. Increase number of quality post-secondary credentials to meet demands of the economy;
4. Better integrate Illinois' educational, research, and innovation assets to meet the economic needs of the state and its regions

These four goals involve using the ILDS in numerous ways and one of the most important ways will be in conducting research based on the longitudinal experiences and outcomes of individuals who have records maintained in the system. This report, developed with a grant provided to DePaul University by IBHE, is intended to help those who are doing the longitudinal studies. It is not exhaustive in its content or discussions. It is intended to address key issues and concerns and to provide pathways to additional material.

There are two parts to this report, the first part looks at using a longitudinal data system. It provides an introduction, context, and examples of longitudinal studies. The second part focuses the technical aspects of a longitudinal data system. It includes construction of a longitudinal data system, methodological concerns, the technology of storage and display, and the different sources of data that may be included. There is an Appendix that includes a glossary of the multitude of acronyms so prevalent in a technical discussion. The Appendix also includes references used in this discussion that provide additional material on the topics discussed.

While this discussion focuses on the longitudinal data system, the parts and their sections are also designed so that readers can go directly to the aspects most important to them. The following is a brief description of the two parts and their sections to aid in focusing on the aspects of the document that are most relevant to specific issues.

After this introduction, the first section of Part 1, (Section 1) provides a working taxonomy to categorize the different types of longitudinal studies that are conducted and their characteristics, comparing them to the primary alternative of cross-sectional research. The second section (Section 2) goes into more detail about the types of questions that may be addressed with longitudinal studies and the advantages of this approach, citing many of the key scholars and practitioners of this research. The third section (Section 3) explores the context for longitudinal databases and their use in relevance to various funding initiatives, national standards, federal reporting, policy research, state oversight, and policy agendas. Examples of longitudinal studies at nation, region, state, and institution levels are the focus of Section 4, with a review of their use in scholarly studies and graduate work. Finally, the last section of Part 1 (Section 5) gives an overview of the different types of benchmarks and performance measures that may be calculated with longitudinal data, including those related to community college student success, transfers, student typologies, and assessment.

The second part of the report focuses on building and using the longitudinal database in longitudinal research. Section 6 discusses the practical aspects of building a longitudinal dataset, including extracting data, data structures, data manipulation issues that arise, data integrity, the use of student identifiers, defining cohorts, crosswalks and taxonomies for categorical variables, and tracking time to completion. Methodological concerns are presented next in Section 7, including design issues, advanced statistics, issues in instrument/survey construction, problems with definitions of "value added," program evaluation, sampling, and the use of multiple methods.

The technology of data storage and display is the topic of Section 8, including current expectations for dynamic web display, incorporation of new business intelligence (BI) tools, open source and free alternatives for software, and observations about the information technology change process. Different sources of data and key variables of interest are highlighted in Section 9, including P-12 (Pre-Kindergarten to High School) schooling, social services, financial aid, employment, industry certifications, workforce training and non-credit instruction, and learning management systems. In conclusion, some planning issues are put forward for consideration by the reader in Section 10. Other issues such as the Family Educational Rights and Privacy Act (FERPA), data-sharing agreements, security, access, project planning, Information Technology support (to include the cloud), Business Intelligence, and data mining and visualization are beyond the scope of this monograph. While incidental mention is given, the reader is referred elsewhere; where possible, however, resources are provided for further exploration.

Throughout the report, the first occurrence of an acronym or abbreviation in the text is spelled out in each section. The Appendix in Section 11 contains a list of abbreviations, references, and a list of websites mentioned.

# Part I: Context, Measures, and Examples

The numerous decisions that are made in designing Student Unit Records (SUR) and building systems for their collection require that we understand longitudinal research and determine the data elements that support the types of decision-making that is done. The purpose of this publication is to help with these needs. This presentation must start with broad questions, such as: What is longitudinal research? What work has already been done on longitudinal research about higher education? What has been learned? What are some of the tools that can be used to study questions with this approach?

The purpose of a longitudinal database will frequently drive the technical design of what data are to be included in the database, how the database is organized, how the database is managed, and who is involved in key decisions concerning the database. The core data in the Illinois longitudinal database are established by the Higher Education Consortium. It is envisioned that data will be integrated with other data across the educational experiences of students. As these data are shared with institutions, it is expected that the institutions will augment the core data with local, contextual data to inform institutional decision-making. As the institution creates its own data mart of student data, it is important that its staff consider the alternatives and uses of the longitudinal data. The following discussions are intended to help with the decisions that an institutions needs to make in selecting and managing its student longitudinal databases.

## Section 1. Types of Longitudinal Studies

Based on the approaches outlined above and this review, a working list of longitudinal study and report types may be developed. These include:

1. National, sample studies (ex. National Center for Education Statistics (NCES), the Beginning Postsecondary Students(BPS) and the Wabash National Study Surveys);
2. National, population studies using the National Student Clearinghouse for special topics (ex. national attainment rates, transfers)
3. National, federal population studies as part of oversight of financial aid (ex. Free Application for Federal Student Aid (FAFSA), TRIO Programs, scholarships, Gainful Employment)
4. Quantitative studies for research (ex. *Research in Higher Education* articles, dissertations, and the work of Pascarella and Terenzini );
5. Analysis of unit record level data collected by state agencies for policy analysis (ex. affect of a new financial aid, grant policy);
6. Institutional studies to understand persistence and completion, program evaluation, assessment,

and student achievement and institutional effectiveness for accreditation purposes;

7. Unit record studies of one or more states/systems conducted by policy organizations, associations, foundations, and others to improve student success or to monitor research grants and scholarships (ex. Achieve the Dream (ATD), Jobs for the Future, Gates);
8. State longitudinal data systems (SLDS) focused on the continuum from P-12 to the workforce;
9. Reports based on categorical characteristics about cohort progression and achieving outcomes;
10. Other purposes not listed above.

While this monograph will provide some support for all of these purposes, its primary focus is with the sixth purpose – using the SLDS in institution studies to examine a variety of questions and interests such as retention, interventions, value-added college impact, and institutional effectiveness expected as part of regional, national, and disciplinary accrediting agencies. Yet the same data may be of use for a dissertation; or a sampling frame may be necessary due to the use of a particular instrument to ensure stratification. It is helpful to think about studies done that are comparable to one's one and to look for effective practices, such as how previous researchers have addressed the vagaries of defining transfer and intent to complete.

## A Taxonomy of Longitudinal Studies

The term "longitudinal study" can imply several different methodological designs and reasons for research. The purpose of this section is to assist researchers in using the data that are collected as part of a specific state longitudinal data system (SLDS) initiative. However, there are many ways in which longitudinal studies may be approached. These approaches may be thought of in terms of: (1) organizational focus; (2) topic/question; (3) respondents; (4) nature of inquiry; (5) quantitative methodology; (6) use of cohorts; (7) time periods examined; (8) granularity/levels of aggregation; (9) use of the research-driven, empirical knowledge base; (10) use of questionnaires, surveys, and instrumentation; (11) mandate for development; (12) use of multiple types data within a general source; (13) use of multiple data sources across the continuum; (14) particular data structure and storage; and (15) technology for dissemination. The following gives additional detail on the types of studies developed under these different approaches and also gives examples of some of the studies that have been done.

| Approach | Types | Examples |
|---|---|---|
| **Organizational focus** | Nation, region, multi-state, state, system, sector, control, institution, program | NCES BPS Longitudinal study, Wabash, SURE, Multi-State "Human Capital Development Data System, NSC Signature Reports, retention committee |
| **Topic/Question** | What is the primary topic or question of study? | Students, faculty, financial aid, expenditures, revenues, publications, research |
| **Respondents** | Entire population, sample, panel | Panels sampled from NPSAS & tracked over time by NCES, GRS subgroups |
| **Nature of inquiry** | Assessment, retention study, theory development, policy question, dissertation/thesis | Use of student engagement theory & surveys |
| **Quantitative methodology** | Multivariate, multiple methods, qualitative, descriptive, data mining | Cox regression, path analysis, factor analysis |
| **Use of cohorts** | Demographic/program breakouts | IPEDS Graduation Rate Survey, IPEDS retention ratios |
| **Time periods examined** | Cross-sectional with different groups at same time, repeated measures, multiple snapshots over time | Beginning vs. graduating students in same Spring semester, longitudinal data by semester over 10 years |
| **Granularity/Aggregation** | Organization, division, program, subgroup, or individual level view of data | Data collected for program intervention over time, broken out by cohorts; view data at program level with interventions |
| **Research knowledge base** | Based on review of cumulative literature & research, foundation agenda, policy question, | Review of Pascarella & Terenzini, Ewell, RHE, AIR Professional File & IR Resources, Tinto, others; Lumina Big Goal, ATD, JFF, NGA, CCA[1] |
| **Use of questionnaires, surveys, & instrumentation** | Vendor, research-driven, home-grown, none | Use HERI, CLA, NSSE, Holland, etc. vs. doing analysis based on existing research & theory |
| **Mandate for development** | SLDS requirement of Federal/ARRA stimulus monies, WIA, state SUR data collection, IPEDS GRS, VFA, VSA | Build to meet funding requirement, as part of grant, to meet association expectations, to meet legislative requirements |

[1] These and other acronyms are explained in the Glossary in the Appendix.

| | | |
|---|---|---|
| **Multiple data types within source** | Ex. in post-sec: admissions/ testing, enrollment, financial aid, course-taking, awards, licensure, learning communities, finance | Bring together different sources for derived, value-added variables of interest, such as attaining milestones, Tipping Points |
| **Linking Data sources across the continuum** | | P-20 to workforce, social services to post-sec, none |
| **Data structure & storage** | Complex relational model requiring query tools, functional tables or views in data warehouse, data mart with single data source | Data warehouse in Oracle with OBIEE tools |
| **Technology of dissemination** | | Visual Tableau, Zogo, Micro-Strategy |

While the national call to action and policy context for accountability may be heard differently at the institution level, most of the issues that must be addressed with longitudinal studies of this nature are the same.

Bauer (2004) presents four types of longitudinal designs, citing Menard (1991). The first type, *total population designs*, involves a study of an entire population over time in different time periods. It is understood that the number of records/cases will change with death, dropout, and other reasons; but group change and trends may be examined. *Cross-sectional designs* are the second type and are the most popular, with samples/cases drawn at one or more times. Groups are included that are at different stages of maturity. For example, when an entire group of students are studied and results are interpreted based on a freshman to senior student level, this is a cross-sectional design. Examples include the College Student Experiences Questionnaire (CSEQ) and the Higher Education Research Institute's (HERI) First Year College Survey. With this design, it is possible to study aggregate trends during one period. However, this can't be used to "study developmental patterns within a cohort and to examine causal relationships" (Bauer, 2004, p. 78). The cross-sectional design is a "'snapshot' of the influences at a single point in time" (Terenzini, 1987, p. 28). However, "Differences found between or among groups may be due to differences between or among the groups at the time they enrolled. Failure to consider such pre-college differences may lead the researcher to conclude (unwittingly and perhaps expensively) that the sources of attrition lie within the institutions control when, in fact, they do not" (p. 28). The strength of the cross-sectional study is that it can be done at a point in time. The limitation of the cross-sectional study is that assumptions must be made that the groups represent sets of similar individuals who differ only on being at different stages of maturity. This methodology is vulnerable to factors that limit its generalizability.

A "sizable percentage" of published research and "the vast number of smaller-scale institutional assessment efforts" use a cross-sectional design to measure the impact of college on learning outcomes. This has led to the "likely inaccurate attributions of curricular and co-curricular programs to institutional quality or effectiveness (whether they are positive or negative)" (Siefert et al, 2010, p. 13).

With longitudinal data disaggregated for student groups, there is "information about where to intervene to improve outcomes," explains Prescott (2011, p. 24). The data "do not tell you *how* to intervene; that requires an inclusive, engaged, and iterative process."

The third type of longitudinal study, *revolving panel designs*, collects data "on a sample of cases for a specified measurement period, then drops some subjects, who are replaced with new subjects" (Bauer, 2004, p. 78). This addresses problems in cohort attrition, but allows examination of individual change.

The fourth type, *longitudinal panels*, studies cases over time, usually with multiple cohorts that "enable analysis of age, period, and cohort effects; description of developmental and historical change; analysis of temporal order events; and causal analysis" (Bauer, 2004, p. 79).

As indicated above, the two main types of studies using student data are cross-sectional versus longitudinal panel. Cross-sectional designs are less expensive, quicker to conduct, and can include more subjects in the design for a given budget. Bauer (2004) explains that "the strengths of the longitudinal panel design, however, are the weakness of cross-sectional design," – the "ability to identify individual variation in growth or to establish causal relationships between variables. Collection of data on individuals at three or more points enables powerful statistical modeling techniques, and the precision with which parameters of growth can be estimated improves with each additional wave of data"( p. 79).

Since longitudinal studies are usually ongoing and evolving, a new instrument or topic may be introduced after other waves of collection have been completed. Terenzini explains that "If a cross-sectional design provides an informational snapshot of the influences on students' attendance behavior at one point in their college careers, longitudinal designs constitute something of a family album" (Terenzini, 1987, p. 29). Subjects in longitudinal studies are exposed to time specific events, where different cohorts will have different experiences. However, the use of repeated measures gives more control over the interpretation of such events and comparisons between cohorts can help in the interpretation of their effect.

The value of longitudinal panel studies is the reason for the growth in Student Unit Record (SUR) data systems[2]. Sampling may or may not be involved, depending upon the use of instrumentation with pre-

[2] There are also many uses for institutional longitudinal data in higher education. For example the Delta Project includes historical IPEDS and other institutional data on finance, affordability, and cost-related variables by institution and can be longitudinal. The data may be rolled up to sector, Carnegie type, state, and other variables of interest.

/post-measures for assessment. The National Forum on Education Statistics' (NFES) Forum Guide to Longitudinal Data Systems explains that "there is no one way to build an LDS and no two such systems are alike. Each educational organization takes its own path to its own version of an LDS, fulfilling its own specific set of goals" (NFES, 2010a, p. 2).

Longitudinal SUR data can improve our understanding of the learning process and help shape policies to improve the educational process. "Policymakers see a growing need for solid longitudinal information about student progression" and "business and civic leaders recognize how vital this 'supply chain' of educational capital is in their states" (Ewell and Boeke, 2007, p. 2). "Only a set of robust longitudinal data on the characteristics and experiences of each student… provides the ability to thoroughly investigate the patterns of success and struggle that students experience" (NFES, 2010a, pp. 8-9).

## The Importance of the Longitudinal Study

While longitudinal studies may appear at first to be daunting and complex, they are a "mainstay" and "at the core" of institutional research (AIR, 2010; Seifert, 2010). "They are among the most valuable work that can be done to examine student success" (AIR, 2010, p. 34). Longitudinal studies "have provided college and university administrators with a plethora of support in making data based decisions related to policies and practices" (Rocconni and Ethington, 2009, p. 368).

While it might seem that longitudinal pre-/post- designs are "widespread in the literature and in practice," these are mostly cross-sectional and true longitudinal studies are not as prominent (Siefert et al, 2010). Researchers such as Pascarella have "called for an increase in use of longitudinal data with pretest-posttest design when studying effects on college students" (Rocconni and Ethington, 2009, p. 368). Unfortunately, "Most institutional research (IR) practitioners lack access to direct measures of longitudinal student learning, let alone assessment instruments systematically embedded into the curriculum" (Herzog, 2011, p. 21).

Both the relatively simple calculation of fall to spring retention rates and the data submitted on the IPEDS Graduation Rate Survey (GRS) require institutions to study students at points in time and report their progress using standard performance measures. Some form of longitudinal study is in place at most institutions, though it may not be labeled as such.

The expansion of this approach with different cohorts, time periods, and variables of interest affords an almost infinite array of possibilities for research and analysis. These possibilities are well understood by policymakers, researchers, and vendors across the P-20 spectrum from preschool to graduate education, many of whom are working at the tail end of a federal funding frenzy to build statewide longitudinal data systems. In approaching the development and use of longitudinal studies, it is imperative that IR professionals understand their context, especially within the current climate of policy discussion in federal and state government and national and regional accrediting agencies about accountability and the stewardship of scarce resources. This monograph is intended to provide a basic, reference for conducting longitudinal studies at the institution level. It is written for the small IR office that does not necessarily have dedicated database professionals. There is a great deal of IR literature about developing retention and graduation models and creating longitudinal tracking systems; for example see Ewell, Parker, and Jones (1988), *Establishing a Longitudinal Student Tracking System: An Implementation Handbook*. The professional development module "Longitudinal Tracking for Institutional Research" (AIR, 2010) is another important resource. Part of the AIR Data and Decisions Academy, this online training module provides an introduction to practical issues and uses a series of tasks and exercises to build competencies in this area. There is no substitute for the prerequisite of understanding previous work in this field.

It is important that IR practitioners understand the kinds of questions that are being addressed with the funding of large, longitudinal data systems. States are tying together K-12, postsecondary, and workforce data in new ways. The institution-level study should not be done in a vacuum. Rather, national, state, and system data need to be leveraged wherever possible, especially for transfer studies. The "ideal state data system" promoted by National Center for Higher Education Management Systems (NCHEMS), State Higher Education Executive Officers (SHEEO), Data Quality Campaign (DQC), Common Education Data Standards (CEDS), State Longitudinal Data Systems (SLDS), and other initiatives is the same ideal system recommended for institutions, albeit one that must be built at the institutional level with far fewer staff and resources.

## SECTION 2. WHY DO LONGITUDINAL STUDIES?

There are a number of reasons for conducting longitudinal studies beyond complying with federal and state mandates and these need to be acknowledged. Foremost is their use in assessment. Terenzini makes it clear that assessment and accountability efforts "need to demonstrate that college and university attendance makes a difference, that students leave colleges and universities with knowledge, skills, attitudes, and values they did not have when they arrived" (2010, p. 38). The following are some of the reasons that various authors have articulated for doing longitudinal research.

A primary use is in supporting enrollment models for enrollment management. In "Principles of Longitudinal Enrollment Analysis, Conducting Retention and Student Flow Studies," Ewell (1987) describes four ways for analysts to support enrollment management with longitudinal data. First, is "to establish a basis for building a model of enrollment dynamics over time;" one that links together key events including admissions, transfer, withdrawal, dismissal, reenrollment, and completion. Second, is to "identify and distinguish the behavior of different kinds of students." Third, as a result of using this model, the user can "estimate the effects of proposed policy changes on total enrollment and on the distribution of enrollment" (p. 2). Fourth, using this longitudinal model will help bring cohesion to the "discrete studies carried out for disparate and particular purposes, the results of which are never used again." It should be considered as a "mechanism for organizing the findings of past research so that they can shed light on enrollment behavior" (p. 3). Ewell's point is well taken. While an important topic for scholarly and episodic research such as dissertations, longitudinal research should be an ongoing and iterative process for institutions; not a one-time project to be completed, presented, and shelved until the next time someone becomes interested in a current policy topic.

The following are some of the authors who have discussed the use and value of longitudinal student research. The intent herein is to give a sufficient description to support the retrieval of the full text of articles by these authors.

Leinbach and Jenkins (2008) argue in *Using Longitudinal Data to Increase Community College Student Success* that "Understanding how students actually progress through their college programs is essential in developing strategies and choosing appropriate interventions to improve student outcomes" (p. 1). These strategies are not always effective because "Many of our institutional and public policies are predicated upon assumptions about college going that are no longer valid" (Ewell et al, 2003, p. i).

The "longitudinal nature of education is implicitly recognized in the extant research literature," note Reynolds et al (2010, p 56). "Longitudinal approaches are essential if we hope to match methods with the phenomena of interest" (Pai et al, 2008, p. 8). With "the ability to track different cohorts of students", it is possible to "define student success measures differently and begin to design financial aid and developmental policies that can enhance" the higher education experience (p. 8).

In his manual entitled "A Toolkit for Community College Data Use," Ewell (2008) documents the use of longitudinal data and provides case studies. Longitudinal data allow us to: examine patterns of progression and completion for particular student populations; evaluate and improve developmental courses and sequences; gauge the effectiveness of career and technical programs; investigate student movement across levels; and estimate future enrollment demand.

Focusing on the use of these data to serve low-income families, Phillips (2009, p. 7) explains that "it is not enough to know whether someone succeeded or failed in a program or in achieving a certificate or degree. State leaders need to know at what point someone failed on the continuum and why; for students who succeed, state leaders need to know more what contributed to that individual's success."

Describing the use of this evidence, there needs to be a "data-driven improvement process that includes the right mix of success indicators, goals, incentives, and technical assistance and program implementation supports" (Jobs for the Future, 2010, p. 4). This is a process of "Moving from Collecting Data for Compliance to Using Data for Continuous Improvement." It requires a "culture shift," away from negative perceptions of accountability to proactively rather than reactively "alter policies, programs and practices" (DQC, 2009, p. 2).

## Types of Questions that May Be Addressed

Many questions can be answered with longitudinal data, some with institution-level studies and others requiring state and multi-state systems. For example, "How effective are remedial courses in preparing students with assessed academic deficiencies for college-level work?" and "How important are particular academic experiences or the attainment of particular enrollment milestones… to student success?" (Ewell (2008, p. 2). "Do students take remedial course sequences in the order we intend?" and "How long does it take first-time, full-time students to earn a degree or certificate?" (AIR, 2010, p. 15). What is the profile of a graduating class in terms of demographics, student status, developmental education, and financial aid? What proportion of students who enrolled in a student success class achieved sophomore status?

Longitudinal data are being used to address workforce questions such as "What is the employment rate of graduates that have some postsecondary education compared with those that have earned a postsecondary credential… and in what industries do they work?" (DQC, 2010, p. 1).

A number of "educational pipeline" issues may be addressed, looking at the overall flow of students and asking what curricular and environment factors affect success in making progress. "What facilitates successful student transitions across specific boundaries" and "How are these transitions different for different types of students?" (Ewell and L'Orange, 2009, p. 1).

## Advantages to Longitudinal Studies

The "greatest advantage of longitudinal studies," according to Bauer (2004, p. 79), is "the ability to identify individual variations in growth or to establish causal relationships between variables." Terenzini (1987) explains that the longitudinal design provides "extensive, planned control of confounding background variables, as well as more precise estimates of the institutional influences on attendance behavior. Such designs are the most internally valid available for studying attrition and afford a measure of confidence in findings and associated conclusions that is not available with other designs" (pp. 29-30).

"Longitudinal designs are the most demanding, but also the most likely to produce valid information" (Terenzini, 1987, p. 29). Still, it must be admitted that "The most difficult and technically challenging question to answer is 'Why do some students withdraw while others continue?'" (Terenzini, 1987, p. 25). "Longitudinal impact studies are essential to any examination of the effects of educational intervention strategies," writes Endo (1992, p. 30). Sometimes, the need is for "a more focused study on a specific intervention strategy (or strategies) and set of outcomes." Alternatively, Terenzini describes how some studies are less focused, instead looking at "the effects of a wide range of loosely defined or unspecified intervention strategies on a wide range of loosely defined or unspecified student outcomes" (p. 26).

## SECTION 3. THE EXTERNAL CONTEXT FOR LONGITUDINAL STUDIES

Longitudinal studies are done within the political, economic, and social milieu of doing more with less resources. The context and knowledge base for longitudinal studies has never been richer or more complex, with significant efforts at the federal and state level, initiatives by foundations and national associations, and advocates in the research and vendor communities. Aspects of these efforts can be grouped into the categories of funding initiatives, standards, national longitudinal studies, national data collections, state student unit record (SUR) systems, and other policy agendas and research efforts. The following sections address each of these categories of effort. There is so much activity occurring in so many different arenas it is impossible to address each activity. It is important however to highlight the major themes of major efforts in order to provide a framework for institutional use of longitudinal data. It is also important to note that the landscape of longitudinal data and their use is very dynamic and detailed use of information should be updated and verified for its completeness and currency. Fortunately, such information is increasingly web-based, provided in dynamic, data-driven applications, not just static reports.

### FUNDING INITIATIVES

A variety of federal initiatives provide funding for states to support longitudinal data collection and improve student outcomes. The Data Quality Campaign (DQC)[3] was founded in 2005 by an umbrella of national organizations and funded by the Bill and Melinda Gates Foundation to improve the collection and use of education data. Tracking different funding and mandates, the DQC developed an interactive roadmap to federal legislation and the work of the U.S. Departments of Education (ED), Health and Human Services, and Labor (DQC, 2012c).

The centerpiece of the federal, postsecondary component is the grant process to create State Longitudinal Data Systems (SLDS). The SLDS initiative was authorized by Congress in the Education Sciences Reform Act (ESRA) and the Educational Technical Assistance Act (ETAA) of 2002. The American Recovery and Reinvestment ACT (ARRA) Stimulus funding in 2009 provided a round of competition with $250 million to expand data systems in 20 states (Gould, 2011). Since November 2005, more than half a billion dollars in SLDS grants have been awarded to almost all states. Three to five year awards have ranged between from $1.5 and $19.7 million. The latest round occurring in summer 2012 with awards to 24 states. The SLDS program is administered by ED's Institute for Education Sciences (IES), which houses the National Center for Education Statistics (NCES). The overall goal of this effort is to "design, develop, and implement SLDSs to *efficiently and accurately manage, analyze, disaggregate, report, and use individual student data"* (Gould, 2011, p. 3).

[3] The reader is referred to the Glossary for acronyms as they are used in many external initiatives.

The legislation for the additional stimulus money that went to states as part of the State Fiscal Stabilization Fund (SFSF) required state governments "to establish P-20 longitudinal data systems and report college enrollment and credit-accumulation rates by state, local education agency, and high school" by student subgroup (DQC, 2012, p. 1). Theoretically, by FY2009, SLDS programs should have included K-12, pre-kindergarten, postsecondary, workforce, and student-teacher data.

Signed into law in 2007, the "Creating Opportunities to Meaningfully Promote Excellence in Technology Education and Science" Act (COMPETES) is another funnel for SLDS activity. It promotes education in science, technology, engineering and math (STEM) fields by agencies such as the National Aeronautics and Space Administration (NASA), National Science Foundation (NSF), and ED. When it was reauthorized in 2010, the America COMPETES Act included 12 data elements that states must include in their SLDS-funded longitudinal data systems.

This list of data elements suggests an inherent design and the type of issues that are being addressed in these systems. These include: (1) a unique identifier; (2) student unit record data on enrollment, demographics, and program; (3) enrollment, transfer, and completion data; (4) K-12 test data required specifically under the EASE act; (5) K-12 student data on those not tested by grade and subject; (6) test scores about college readiness; (7) links to teacher data; (8) course transcripts with grades; (9) college success, including participation in remediation; (10) data on K-12 preparation for college success; (11) data quality audits; and (12) data-sharing from preschool through postsecondary education (ED, 2009).

The Workforce Data Quality Initiative (WDQI) focuses on integrating education and workforce data. Other funding streams such as Race to the Top, the Individuals with Disability Education Act (IDEA), Perkins IV, Title I, the Teacher Incentive Fund, and the Workforce Investment Act (WIA) include language that encourages longitudinal data systems. Race to the Top promotes states building and using longitudinal systems. The competition pushes states to "build the tools to measure results, guide decision-making and investments, and provoke honest conversations about whether and how schools are on track for, meeting, or exceeding college and career readiness goals" (Achieve, 2012).

A Public Domain Clearinghouse is being created to share tools, documents, and effective practices of states in developing their state longitudinal data systems. This is intended to lighten the burden and lower the cost of developing tools, as well as to promote collaboration (Sellers, 2011). This kind of sharing of effective practices is needed because "The intensity of simultaneous activities in this arena may result in efforts that are hurried and uncoordinated, with states independently designing and implementing their own systems. An unfortunate end result may be a patchwork of systems that cannot be easily aligned within a state or across borders" (Prescott and Ewell, 2009, p. 1).

The process of modeling a SLDS is the topic of a dissertation by Olsen (2010). This dissertation reviews barriers to success of state LDS initiatives and focuses on "repurposing business data warehouse systems" for educational research. A vision for a relational model of SLDS is shared, including an interface and visualization tools.

## STANDARDS

The Postsecondary Electronic Standards Council (PESC) promotes data exchange standards, envisioning national interoperability and the seamless flow of data between systems. Included in its mission is "data alignment across disparate systems and across sectors," setting standards, and supporting "business models for data standardization, access, and exchange" (PESC, 2012). Current standards for data elements must be addressed, regardless of the scope of the longitudinal database, and practitioners are advised to keep up with how well their data dictionaries adhere to them.

Common Education Data Standards (CEDS) have been developed in another NCES initiative with funding by multiple foundations. CEDS was created with the recognition that "answering critical questions requires appropriate data to flow efficiently and effectively across systems, sectors, and states" (DQC, 2012). It builds upon previous efforts, such as the CHESS by NCHEMS (Thomas, 2004a). A common language is promoted with "commonly agreed upon names, definitions, option sets, and technical specifications for a given selection of data elements" (DQC, 2012, p. 8).

The CEDS tools allow users to see how their data fit within standard structures, table relationships, variable names, and formats for typical longitudinal reporting. Any variation from expectations helps anticipate problems and save time and resources.

The CEDS Domain Entity Schema is a "hierarchy of domains, entities, attribute categories, and attributes" useful in searching, mapping, and organizing data elements. A Normalized Data Schema or logical data model is promoted as the CEDS framework for P-20 longitudinal data systems. An online tool is provided for users to compare their data element dictionaries to the CEDS (CEDS, 2012). While the CEDS has multiple purposes, like the work of PESC, the current standards and value labels of the data structures should be understood and incorporated when building a system. They are not intended, however, to be prescriptive, but as "a means to have a common vocabulary so that we all speak the same language" (L'Orange, 2012, n.p.). Version 3 of CEDS is to be released in January, 2013.

The Common Core State Standards Initiative is a comparable initiative for K-12 education. Led by the National Governors Association's Center for Best Practices and the Council of Chief State School Officers (CCSSO), the core standards define the knowledge and skills needed in high school to success in college and workforce training (L'Orange, 2012).

The Ed-Fi initiative, which is funded by the Michael and Susan Dell Foundation, expands this approach into the K-12 data world with another national standard that is aligned with CEDS. This initiative promotes tools for building BI dashboards that bring data together from different schools and systems and displays information to various constituents such as teachers and students in new and dynamic ways. Ed-Fi has developed and documented longitudinal, K-12 data elements and many of these are incorporated in CEDS.

The National Forum on Education Statistics (NFES) is an activity of ED's NCES Cooperative System that promotes effective practices in producing and maintaining early childhood and K-12 education data. Four monographs have been prepared as "The Forum Guide to Longitudinal Data Systems." While geared to a P-12 audience, many of the same issues must be addressed by states and institutions building databases for longitudinal studies that span the continuum of lifelong learning.

The first volume, *What is an LDS?*, introduces the longitudinal data system and its concepts and benefits; where the second, *Planning and Developing an LDS*, focuses on stakeholders, needs assessment, procurement, evaluation, and change management. *Managing the LDS* is the topic of the third volume, addressing data governance, data quality, standards, security, and the protection of confidential data. The final volume, *Advanced LDS Usage*, describes ways to use an LDS and how results may be leveraged. Strategies for training and development are also presented.

Much of the material in this series addresses process and communication issues, such as how to identify different needs for the display and presentation of data, and developing strategies for organizational change. For example, sample memorandum of understanding are provided for sharing data between K-12 and social service agencies. These same issues must also be addressed by postsecondary institutions.

## Federal IPEDS Graduation Rate Survey (GRS) Reporting

The IPEDS Graduation Rate Survey (GRS) was developed in response to the Student Right to Know and Campus Security Act of 1990 (SRK). It requires institutions to report "the rate at which full-time, first-time, degree/certificate-seeking undergraduate students complete their academic programs" (IPEDS TRP, 2012, p. 2). Data are collected on a cohort of students and their completion rates at 100% and 150% of time. Under the Higher Education Opportunity Act (HEOA) of 2008, the GRS time period was extended to include 200% of time (for example, 8 years for a four-year bachelor's degree). IPEDS also collects data on transfers in the GRS. The related measures of full- and part-time student retention are collected in the fall enrollment (EF) survey component.

Achieving the Dream (ATD) is a national initiative conceived in 2004 by Lumina Foundation for Education in conjunction with several national associations, policy organizations, and educational research programs. ATD provides a broad umbrella of research and practice efforts designed to "help more community college students succeed." One ATD study collected six years of state, community college student unit records from Connecticut, Florida, North Carolina, Ohio, Texas, and Virginia. The results were analyzed and compared to what can be learned with IPEDS. The conclusion is that the current IPEDS GRS focus on first-time, full-time students is too simple and not a good measure for community colleges, resulting in "rates that are all but meaningless" (Ashburn, 2008).

It was recommended that the GRS be expanded to include part-time, first-time freshmen and that the length of time that students are tracked be expanded to 300%. This would allow for up to six years for associate degree attainment. Also, transfers to four-year institutions should be monitored as a key indicator of success and all sources of statewide enrollment should be included, not just the initial school (Jobs for the Future, 2008). In "Complete to Compete: The 2010-11 National Governors Association (NGA) Chair's Initiative," Reindl (2010) critiques the IPEDS GRS. He argues that there are no graduation rates for part-time, transfer, low-income, or remedial students and that these data are needed by policymakers.

In a classic piece of IR literature, Astin (1997) asks "How good is your institution's retention rate?" and proceeds to critique the student right to know methodology for failing to differentiate between performance versus effectiveness. Institutions should be comparing "actual outcome measures with expected measures based on student input characteristics." Institutions that "look good in absolute terms are actually underperforming in relation to their student input characteristics, whereas others with mediocre rates are actually performing at a substantially higher level than one would expect from their student input characteristics" (p. 656).

HEOA also requires institutions to disclose completion rates disaggregated by subgroups based on gender, race/ethnicity, Pell recipients, Federal Family Education Loan (FFEL) loans without Pell, and those not receiving Pell or FFEL. This is not being done currently by most institutions except in sample surveys and longitudinal studies. NCES is working through its National Postsecondary Cooperative (NPEC) to provide guidance about how colleges and universities can provide these data (NCES, 2010).

HEOA also required ED to create a Committee on Measures of Student Success (CMSS) to help improve two-year graduation rate information. The Committee was charged with helping ED assist institutions in this reporting and the final version of its report was released in December, 2011. It is noted by the Committee that the "Federal [GRS] rates do not take into account students' college readiness and enrollment in remedial coursework, which may delay their progress toward a degree" (CMSS, 2011, p. 4). The Committee report recommends that the GRS include cohorts of part-time, college-readiness, and financial aid breakouts – each disaggregated by race/ethnicity and gender.

The "Action Plan for Improving Measures of Postsecondary Student Success" was released by ED in April, 2012 in response to the Committee report. It states that a revised IPEDS GRS will be developed to "broaden the coverage of student graduation data to reflect the diverse student populations at two-year institutions." As a result, graduation and transfer rates collected in IPEDS are expected to eventually include part-time, degree/certificate-seeking students, as well as "adding non-first-time, degree/certificate-seeking students" (ED, 2012, p. 1).

The expansion of the GRS to include part-time students was addressed by an IPEDS Technical Review Panel (TRP) in 2008, along with the use of a Pell grant cohort. There was consensus among participants in both the 2008 and the 2012 TRPs that the GRS should be expanded to include a part-time cohort; but "it is first necessary to consider how" they are identified (IPEDS TRP, 2008; IPEDS TRP, 2012).

It is difficult to define part-time students based on "a minimum credit threshold," since part-time students may still receive financial aid. Therefore, it was determined that the part-time cohort be defined using the current IPEDS definition for degree/certificate-seeking students (IPEDS TRP, 2012).

The JCAR Technical Conventions Manual has several definitions for part-time students, including extended load and partial load. The extended load student is one: "who, on average, attempts a course load per term that is not enough to lead to graduation by catalog award time, but is enough to graduate by the extended award time (150 percent of catalog award time or normal time to degree, according to federal Student Right-to-Know, that is: more than two years but within three years for a two-year degree; more than four years but within six years for a four-year degree" (JCAR, 1996, p. 26). A "partial load student" is one who "attempts a course-load per term that is not enough to lead to graduation by the extended award time" (p. 26).

The JCAR report sponsored by AASCU, AACC, and NSULGC in 1996 presaged many of the current developments in longitudinal tracking. AASCU explains that the effort "went beyond GRS in several dimensions," suggesting that the definition of success should include students who have graduated, transferred, or continued to enroll (2006, p. 6). Multiple time periods for tracking student success and multiple cohorts for all first-time, transfer, and first-time full-time students are recommended.

The Committee on Measures of Student Success (CMSS) report explains that "An ideal solution to address the incompleteness of data on student progression, transfer, and completion is a coordinated, public, and privacy-protected student unit record system that includes all institutions that participate in Title IV federal financial aid programs… and that covers student enrollment in all states." "The creation of a student unit record system by the federal government is currently prohibited by the HEOA. While efforts are underway to link state data systems, these efforts are uneven, and progress has been slow" (CMSS, 2011, p. 21).

The IPEDS Student Unit Record Feasibility Study conducted by NCES concluded that "a UR system could be done at most institutions given time for implementation" (Cunningham and Milam, 2005, p. xii). While the ED Action Plan did not move forward with the Committee's recommendation to implement this national system of student unit records, CMSS Chair Thomas Bailey stated that "If we really want to know what is happening with our students… we need to track them across institutions in a longitudinal way" (Gozalez, 2012, n.p.).

## Federal Gainful Employment Reporting

Title IV financial aid requires that an educational program lead to a degree or prepare students for "gainful employment in a recognized occupation" (National Student Loan Data System, 2011, p. 2). All non-degree programs and "virtually all" programs at proprietary schools are considered to fall under Gainful Employment requirements. Under the regulatory power of federal financial aid, ED began in 2011 to require institutions to disclose annually a series of GE rates, including completion, average loan debt, loan default rates, and rates of employment in the field of study. In addition, a national student unit record data collection was put in place to collect data from all institutions for all students in GE programs, "regardless of whether or not a student received Title IV student aid" (p. 3).

The student unit record list of GE data elements included one record per student per program per institution for the reporting award year. There were 28 fields of data. In addition to student and institutional identification, these records include a flag for GE programs; program name; CIP Code; credential level; medical/dental residency flag; FFEL/Direct Loan flag; program start date; award year program start date; attendance status as completed, withdrew, or enrolled; program attendance end date; private loans amount; institutional financing amount; enrollment in another program; CIP code of other program; credential level of other program; program name of other program; GE flag of other program; and institutional identifier and name of other program institution. These unit record data are collected for each year. In 2011, five years of data were required to be submitted. A sixth year of data for 2011-12 was originally due October 1, 2012.

In July, 2012, the GE collection was halted when the U.S. District Court vacated several provisions in the regulations. Bills were also submitted in the U.S. House of Representatives to specifically forbid ED from doing this GE collection under HEOA. As a result of these actions, the SUR data collection ceased. However, the "Court did not vacate the GE disclosure regulations at 34 CFR 668.6(b). Those requirements are still in place. The Department has not, to date, provided the GE disclosure form (template) referred to in the regulations… because we are waiting for the Court to rule on the Department's request to reinstate the GE reporting requirements" (Bergeron, 2012a, p. 1). Institutions are still required to update GE program disclosures for programs. These include occupations that the program prepares students for, normal time to completion, tuition and fees, books, room and board if applicable, placement rates where required by states or accrediting agencies, and median loan debt in Title IV, private, and institutional loans.

FSA has used these federal SUR data to calculate what it calls "GE Informational Rates." These include "debt-to-earnings annual rate, debt-to-earnings discretionary rate, and repayment rate, as well as loan medians for disclosures" (FSA, 2012, p. 1). Three loan medians are provided: Title IV loan debt, private loan debt, and institutional financing debt. Lists of programs failing to meet standards over at least three years' time were documented and sent to institutional financial aid respondents responsible for the NSLDS. Reports from the five years of longitudinal data were disseminated publicly and presented at national meetings such as the AIR Forum in 2012 (Bergeron, 2012b).

These five years of historical data can easily be combined using SSN to produce a longitudinal tracking system of use to institutions, states, and researchers. Studies about individual student progress in programs, enrollment, awards, and financial aid could be prepared by FSA and by submitting institutions. An extensive array of value-added variables may be created. The data could potentially be linked to other data that institutions have at their disposal, such as those collected as part of the FASFA. The many policy purposes for these national GE SUR data, like those proposed in the IPEDS Student Unit Record Feasibility Study (Cunningham and Milam, 2005), are one reason the Association of Private Sector Colleges and Universities (APSCU) and others have fought these initiatives.

## National Student Clearinghouse

The U.S. Department of Education collects SUR-level data as part of processing FASFA forms and distributing student financial aid, much as the Internal Revenue Service collects individual taxpayer data and processes payments and refunds with the U.S. Treasury. ED is precluded from using these data for reasons outside of their intended purpose. IRS data are brought into the FASFA process as part of the streamlined, web form, but behind a firewall, and only for verifying earned family contribution. The ED Gainful Employment data collection discussed previously represents a significant expansion of this power. In the interim and for many years, there has been only one, national SUR data collection in place – the National Student Clearinghouse or NSC (Ruddock, 2012).

Created originally in 1993 to document student enrollment as the National Student Loan Clearinghouse, the NSC "represents a different kind of data system, with a potential not fully realized" (Rice and Russell, 2012, p. 242). Currently, 94% of all student enrollment nationwide and 80% of four-year degrees are included in the Clearinghouse, with data submitted by 3,400 institutions. These data are used in the NSC's EnrollmentVerify and DegreeVerify services as a "trusted and authorized agent" for automated student enrollment and degree verification (Shapiro et al, 2012). Over 1.7 million enrollment verifications are done each year through EnrollmentVerify and two million degree verifications are confirmed with DegreeVerify. Nearly 2.5 million students per year use the free Student Self-Service. All student loan guarantors and most student loan lenders and servicers make use of the NSC. Over 1,200+ high school districts and 4,200 high schools also participate. A secondary education research and reporting system launched in 2009 (NSC, 2012).

The NSC has produced a number of reports and analyses over the past 20 years and gives data back at cost to participating institutions through StudentTracker, while adhering to FERPA limitations. Data are available for reporting through the Voluntary System of Accountability. The

NSC Research Center was created in 2010 and one of its most visible products is the Signature Report series mentioned throughout this monograph. A series of Snapshot Reports present data on mobility, persistence, concurrent enrollment, adult learners, interstate mobility, and two-year contributions (NSC, 2011a, 2011b). Many of these topics are reported in the latest Signature Report 4, "Completing College: A National View of Student Attainment Rates" and address the U.S. college completion agenda (Shapiro et al, 2012).

The NSC "enables the nationwide effort to use accurate longitudinal data outcomes reporting to make better informed educational policy decisions leading to improved student outcomes" (NSC, 2012, n.p). With a "near-census national coverage of enrollments and awarded degrees," it is possible to study cohorts of first-time students over time with a number of student characteristic variables, such as age and institution of first enrollment. Demographic data are incomplete and cannot be used in this manner. Most interesting are the NSC examples of derived variables about enrollment behavior and completion that can be created, such as enrollment intensity, and the enhanced understanding of mobility.

Shapiro et al (2012) define a cohort of 1,878,484 first-time-in-college, degree-seeking students in fall 2006 and track them for six years. The results of measures for first completion/award, persistence, and stop-out are reported by student age at entry, enrollment intensity, and type of starting institution. Attendance is analyzed with a threefold typology of exclusively full-time, exclusively part-time, and mixed enrollment. Completions out of state are also studied, as are four-year completions relative to prior associate's degree completion.

This kind of longitudinal, national portrait is necessary because "The complexity of the postsecondary pathways of today's students makes serious engagement with college completion difficult when using traditional inquiry approaches" (Shapiro et al, 2012, p. 13). A number of interesting patterns of enrollment behavior and completion are shown. Most striking is that the overall success of students is under-reported by the traditional IPEDS GRS cohorts, once all sources of continued enrollment and completion at taken into account. The Signature Report series provide interesting derived variables to include in SLDS databases and provide national benchmarks for comparisons.

## National Associations

The Voluntary Framework of Accountability (VFA) is an effort of the American Association of Community Colleges (AACC) to develop a comprehensive system of metrics for student outcomes and success. Funded by the Lumina and Gates foundations and in partnership with the College Board, VFA was built because IPEDS GRS rates are "not the right tool for measuring community college success" and do not capture the value of the "full range of community college offerings" (AACC, 2012).

A variety of student longitudinal tracking efforts are specified in the VFA, including developmental progress measures such as course completion; two-year progress measures such as credit hours completed with thresholds and persistence; six-year progress measures such as awards, transfers, and continued enrollment; and workforce, economic, and community development measures such as career and technical education (CTE) awards, licensure, noncredit course activity, and Adult Basic Education/General Education Development (ABE/GED) completion. Note that the VFA has gone through a pilot study and its organizers are evaluating the utility and feasibility of some of these measures, including their reporting burden.

A Voluntary System of Accountability (VSA) was initiated in 2007 at the public, four-year level by the Association of Public and Land-grant Universities (APLU) and the Association of State Colleges and Universities (AASCU). The VSA objectives are to demonstrate accountability and transparency, create a college search tool, and support the measurement and data reporting of student learning outcomes (SLOs). A pilot period was completed in December, 2012 and evaluated by the National Institute for Learning Outcomes Assessment (NILOA) with focus groups, interviews, surveys, and data analysis (Keller and Gore, 2012).

The VSA includes a Success and Progress Rate (S&P) as a new metric to gauge student progress. It is recognized that a "majority of students now attend more than one institution before they graduate." The VSA recommends that S&P rates be calculated with data from the National Student Clearinghouse (NSC) for two cohorts: (1) first-time, full-time students (same as the IPEDS GRS); and (2) full-time transfer students. No part-time student rates are included.

The VSA recommends that assessment be done longitudinally at entrance and exit points for these two cohorts. A standardized instrument needs to be used to measure student outcomes and academic progress. Three particular instruments are mentioned: (1) the critical thinking and writing essay modules of the Collegiate Assessment of Academic Proficiency (CAAP); (2) the Collegiate Learning Assessment (CLA); and (3) the critical thinking and written communication sub scores of the Measure of Academic Proficiency and Progress (MAPP).

"Value-added" is calculated in the VSA, based on the methodology promulgated by the Council for Aid to Education (CAE) for the CLA instrument. VFA methodologies and criteria for value-added are documented for each instrument. Typically, this involves comparing within-school differences. A more recent approach is "comparing the CLA performance of seniors at one school to the CLA performance of seniors at other schools admitting students with similar academic skills" (Steedle, 2009, p. 3).

Results from the pilot study of VSA show that almost half of the participating institutions have not met the expectations of the pilot in regards to posting information about student learning outcomes. This section of the College Portraits search engine garners few viewers too, it is noted. It is found that "The standardized test measures of student learning outcomes lack broad credibility and acceptance in the higher education community, undermining institutional participation and engagement with the VSA and campus faculty and staff support of the VSA initiative" (Keller and Gore, 2012, n.p.). As a result, the range of assessment tools and approaches is being expanded and there is more "focus on specific audiences and communicating meaningful information."

In evaluating the VSA pilot, it is recognized that "no perfect measure of student learning outcomes exists for all audiences." While the use of the CAAP, CLA, and ETS Proficiency Profile is continued, rubrics about written communication and critical thinking from "VALUE Rubrics," based on work of the American Association of Colleges and Universities (AAC&U), are being introduced. There has been discussion about schools using aggregate data from graduate and professional admissions exams such as the GRE, GMAT, LSAT, and MCAT. The GRE General Test is currently listed as one of the instrument options. Reporting options now include both value-added and benchmarking.

Neither the VFA nor the VSA provide guidance on how to construct the longitudinal database needed to report the measures they recommend. However, the nature of the data elements themselves implies an inherent structure that is quite comprehensive, with multiple types of cohorts with subpopulation breakouts that are tracked over time.

The private college equivalent to the VSA and VFA was developed by the National Association of Independent Colleges and Universities (NAICU) and is called the University and College Accountability Network (U-CAN). U-CAN does not include learning outcomes data per se. NAICU explains that "extensive research in shaping U-CAN found no consumer demand for learning outcomes data" and that "there is no one learning outcomes measure - or one set of measures - that are broad-based enough to be used across all types of institutions and all academic fields of study." NAICU promotes other initiatives on student learning outcomes and the association and its members work with the National Institute for Learning Outcomes Assessment on projects such as "Making Student Learning Evidence Transparent" (Jankowski and Provezis, 2011).

A College Scorecard was proposed by the White House in February, 2012 to provide a similar set of metrics to those presented by the VFA, VSA, and U-CAN initiatives. In addition to costs and debt, it includes metrics that require complex longitudinal data, including how likely students are to graduate, how long it will take to graduate, and earnings potential. There are a number of problems inherent in collecting and presenting these data. Carlos Santiago, CEO of the Hispanic College Fund, noted for example that "Earnings potential seems to be an impossible measure to truly capture. Earnings differences across disciplines, while available, provide little information for a student's immediate employment/earnings prospects" (Elfman, 2012, n.p.).

## State Student Unit Record (SUR) Systems

In promoting the development of longitudinal data over many years, Ewell has documented data elements and the challenges involved in gathering, linking, and using data. In *Following the Mobile Student: Can We Develop the Capacity for a Comprehensive Database to Assess Student Progression?* (Ewell et al, 2003), he systematically analyzes the capacity of states to track students through SURs, portending the emergence of SLDS. It is noted that a "substantial portion of the nation's enrollment" is captured in SUR systems.

This work was updated with a second, in-depth review (Ewell and Boeke, 2007), with further comparison of data elements, definitions, and coverage. Historical data, periodicity, identifiers, and multiple types of data are documented. "Definitions and coding structures among all of these data elements are sufficiently compatible that they can be linked" (p. 3). Longitudinal data are discussed in terms of how they may be used to understand graduation and retention, transfer, concurrent enrollment, job placement, high school feedback, developmental education, and distance learning.

Additional work about developing state SUR data systems was conducted by the State Higher Education Executive Officers association (SHEEO). SHEEO held a series of national dialogues with data experts and developed a list of "15 essential characteristics," along with recommendations for the "ideal state system" (Ewell and L'Orange, 2009).

Most SURs collect four categories of data: student, course, "operational characteristics," and "data governance" (Ewell and L'Orange, 2009, p. 2). The SHEEO ideal SUR system should also include financial aid, FERPA privacy protections, and data from private institutions. At a minimum, and based on Ewell et al's (2007) documentation of available data elements, the ideal system should include basic student demographics, institutional characteristics, student academic background, current enrollment status, financial aid status, academic activity, and academic attainment (L'Orange, 2009).

The 15 "essential characteristics" of the SHEEO SUR system and the "Ten Essential Elements" of SLDS promoted by the Data Quality Campaign (Ewell and L'Orange, 2009) are very similar to the ten requirements specified under the COMPETES Act of 2007 (Ruddock, 2012). All of these appear to have made their way into the 12 data features required in SLDS under American Recovery and Reinvestment Act (ARRA) and State Fiscal Stabilization Fund (SFSF). These evolved from Ewell's work, were grounded in SHEEO perspectives, were promoted as effective practices by the DQC, and then became requirements of states' accepting federal stimulus money under COMPETES, ARRA, and SFSF.

## Policy Agendas

The Data Quality Campaign tracked the capacity of states to link P-20/Workforce data systems up until 2011, when it was recognized that these efforts were reaching fruition. "This means that, without exception, every state in the country has robust longitudinal data that extend beyond test scores and could inform today's toughest education decisions" (DQC, 2011, p. 1). However, as of 2010, only 10 states were able to link data across the entire continuum from K-12 to postsecondary education to workforce.

Funded by the Gates, Carnegie, Ford, Lumina, and Kellogg foundations, the "Complete College America" initiative (CCA) is a byproduct of a National Governors Association effort. CCA helps document state-level progress toward the goal of increasing degree completion, compiling data across institutions and states from numerous sources. A number of projects are designed to work together to promote completion, ranging from improving gatekeeper course success in developmental math to promoting models for accelerated learning.

One of the guiding principles of Achieving the Dream (ATD) that has emerged is that data should be used as part of a "culture of evidence. " It is expected that the "ideal" ATD college mines student transcripts and other data to understand how students are performing over time — for example, whether students are progressing through developmental courses and into college-level courses, and whether students are returning in subsequent semesters. Colleges are also expected to analyze data by race or ethnicity, income, and other background characteristics to identify groups that may be in need of extra support or intervention" (Rutschow et al, 2011, p. 10).

A data-driven improvement process using longitudinal studies and focused on developmental education is put forward by ATD, Jobs for the Future (JFF), and others. The goal is to "accelerate the creation and scale of new solutions that dramatically improve outcomes for students who test into developmental education" (JFF, 2010, p. 3). This requires that states, systems, and institutions "strengthen their longitudinal student data systems in ways that enable them to track outcomes and reveal which institutions are getting better results… with the right mix of success indicators, goals, incentives, and technical assistance and program implementation supports" (p. 4).

In "A Stronger Nation through Higher Education," Lumina Foundation for Education put forward its "Big Goal" for 60% of Americans to earn a postsecondary credential by the year 2025. An annual series of progress reports is being released. President Obama has focused national attention on educational attainment, calling for increased completion as an "economic imperative." Numerous national higher education associations and states have set their own goals for increases in degree production, often tied to the creation and use of longitudinal data systems. These have been followed by attainment goals at colleges and universities (Lumina Foundation, 2012).

Yet, as Brian Prescott (2011) of the Western Interstate Commission for Higher Education (WICHE) asked in a presentation to SHEEO/NCES, "We have goals… now what?" Student progress and achievement are necessary to unpacking the goal. Using data from the Nevada System of Higher Education, he shows how tracking milestones can lead to better understanding of outcomes and goal setting and benchmarking by subgroups, such as Hispanic/Latino students.

Programmatic changes should be selected that will have the greatest impact and these have been labeled "Intervention Zones." Longitudinal tracking for program evaluation of this kind includes the following steps: (1) define the cohort and key milestones for longitudinal analysis; (2) calculate success rates disaggregated by race/ethnicity; (3) target something for inquiry and action; (4) document success rates by milestone for disaggregated student groups; and (5) monitor progress of goals for equity and completion (Prescott, 2011).

The "New Leadership Alliance for Student Learning and Accountability" is another national initiative developed in response to the call for increased college attainment. By 2018, the U.S. will have "several million fewer degree recipients than the economy needs" and "closing this gap" requires that "colleges and universities enroll and graduate students from previously under-represented populations, including minority group, first-generation, and non-traditional age students" (New Leadership Alliance, 2012, p. 2). In the Alliance report "Committing to Quality: Guidelines for Assessment and Accountability in Higher Education," the questions are asked "Is Your Institution Setting Ambitious Goals?" "Is Your Institution Gathering Evidence of Student Learning?" and "Is Your Institution Reporting Evidence and Results?" The only way to document progress is with evidence from longitudinal studies that is disaggregated for under-represented student groups. The Alliance is endorsed by numerous associations and is linked to NGA, CCA, SLDS, ATD, DQC, and other initiatives. This is another sign of the expanding national attention being given to longitudinal studies and evaluating the results for improvement.

## Section 4. Examples of Longitudinal Studies

While the focus of this publication is on conducting institution-level longitudinal studies, it is helpful to be aware of national and regional efforts that pave the way and provide effective practices. The following section looks at national studies. These are primarily the longitudinal panel studies done on national samples by NCES. In addition to these are several regional studies, often by states that link their SUR state level systems together. There are state-specific studies where SUR data are used to inform the state policy agenda. There are many examples of institution-level studies and several are highlighted.

### Federal Longitudinal Panel Sample Studies

NCES has conducted multiple sample-based, longitudinal, panel studies. The Baccalaureate and Beyond (B&B) and Beginning Postsecondary Students Longitudinal Study (BPS) track student cohorts that were identified using the National Postsecondary Student Aid Study (NPSAS). The B&B built upon the Recent College Graduates (RCG) sample survey that was conducted regularly between 1976 and 1991. Two cohorts of the B&B and three of the BPS have been tracked so far. Both include questions about educational experiences and employment.

High School and Beyond (HS&B) includes two cohorts of sophomore and senior high school students in 1980, tracking them over time with questionnaires and interviews, as well as information from parents, teachers, high school and postsecondary transcripts, and financial aid through 1992. The National Longitudinal Study of the High School Class of 1972 (NLS-72) includes data on high school, postsecondary education, and the workplace from 1972 through 1986 and includes postsecondary transcripts. Tracking a cohort of eighth-grade students in 1988 through 2000, the National Education Longitudinal Study of 1988 (NELS-88) focuses on critical transitions in education and careers. The Education Longitudinal Study of 2002 (ELS:2002) follows a cohort of tenth grade students through high school, postsecondary education, and work. The High School Longitudinal Study of 2009 (HSLS:09) follows a cohort of ninth grader students from Fall 2009, with a second wave of data in 2012.

Longitudinal, undergraduate transcripts from these different studies are brought together in the Postsecondary Education Transcript Collections, providing data about student course-taking, transfer credits, and outcomes (NCES, 2012). Adelman and others have made exhaustive use of course transcripts to understand the nature of undergraduate education. The Postsecondary Education Transcript Study (PETS) brings together transcripts from BPS 2004 and 2009 participants. Using these data, it is possible to study "momentum points and milestone metrics for a national community college cohort that has been followed for six years" (Completion Arch, 2012, p. 67). Items are available in PETS, for example, that ask whether test prep courses were taken for the ACT and SAT and these may be used to study college preparation (Ishitani and Snider, 2006).

The NELS provides a "rich source of data for research on college access, choice, and persistence during the early years of college," explains Berkner (2000, p. 103), who calls it "a tantalizing treasure trove of data" (p. 105). NELS includes data on demographics, socioeconomic background, academic background, educational aspirations and plan, financial aid, concerns about college costs, and information about college choice. Berkner looks at the transition to college and provides an introduction to the NCES longitudinal study program. He describes the multiple layers of data, CD releases of public use data, the Data Analysis System software, the electronic codebook and data dictionary, as well as issues of weighting. The technology for deploying and dissemination these national sample survey data has evolved since 2000 with new online data tools.

Ewell (2008) uses the BPS and NELS to create benchmarks for community colleges. Looking at the percentage of community college students that complete the associate's degree in a period longer than three years, he uses BPS data that suggest that 39% of students attain some kind of credential within six years and another 12% do not get a credential but transfer to a four-year college. Data from NELS shows that 50% of first-time community college students earn some type of credential within 6 to 8 years, with another 13% that do not get a credential but transfer to a four-year college.

Some other interesting examples of using the federal, longitudinal sample surveys for empirical research are Ishitani and Snider's (2006) study of the "Longitudinal Effects of College Preparation Programs on College Retention" using NELS: 1988-2000 and PETS; Walpole's (2008) study of how social class is related to the college experiences and outcomes of African American students; O'Toole et al's (2003) longitudinal analysis of part-time student enrollment and persistence using the BPS 1990/1994; Engberg and Wolniak's (2010) work using ELS: 2002 on how high school context affects students' postsecondary outcomes; and Price's (2004) analysis of B&B data to examine the "relationship between educational debt burden and student race, ethnic, gender and income characteristics" (p. 701).

The tools available to researchers for accessing and using these federal datasets have evolved over time. They have been released in public use files, in CD software packages, in online and PC-based software called Data Analysis Systems (DAS), and most recently as part of a larger PowerStats application. See PowerStats at http://nces.ed.gov/datalab/. Users may obtain licenses to use the student level micro data. In each iteration of these software efforts, extensive codebooks and data element dictionaries have been shared. Websites such as the NCES/NPEC ANSWERS project have been developed over time to explain how the data may be used, with links to data analyses and publications (Milam, 2003). An institution that has longitudinal data and has merged it with additional institutional data might well consider creating a data mart and putting the data mart together with an analytic tool and creating the ability to look at various research and policy questions in a manner similar to that supported by Datalab.

Graduate level data on the entire population of doctoral recipients are available in the Survey of Earned Doctorates (SED) conducted by the National Science Foundation (NSF). These are linked to longitudinal panels in the sample-based Survey of Doctorate Recipients (SDR) conducted by NSF. The NSF website has extensive information about these and other graduate level sample surveys, but these are the only two that are longitudinal in nature. Unfortunately, the number of SED respondents sampled in the SDR precludes much longitudinal analysis. Online and CD-based data tools such as CASPAR, WebCASPAR, and SESTAT have evolved over time at NSF like at NCES. These tools make population and sample survey data available about graduate students, doctorate recipients, and scientists. The data are available at the individual level with site licenses (Fink and Muntz, 2012; King et al, 2012; Milam, 2003).

## Regional Studies

The Wabash National Study of Liberal Arts Education is a prominent longitudinal study, considered "one of the largest longitudinal panel studies undertaken in the United States" (Siefert, 2010, p. 2). It was funded by the Center for Inquiry in the Liberal Arts at Wabash College and is "a large, pretest/posttest longitudinal investigation of the effects of liberal arts colleges and liberal arts experiences on the cognitive and psychosocial outcomes theoretically associated with liberal arts education" (Padgett et al, 2012, p. 247).

Five chapters in a special *New Directions for Institutional Research* Assessment Supplement about longitudinal studies are based on the Wabash Study. A variety of institutions with a range of size, selectivity, and location participated and cohorts of students starting college in fall 2006, 2007, and 2008 are tracked. Two cognitive and five affective sets of measures are used (Herzog, 2011).

Padgett et al provide an example of how the Wabash data may be used with pre-/post-testing to investigate "the effects of college experiences on a range of liberal arts outcomes" (2010, p. 34). A pre-college survey was used that included demographic and background variables, as well as measures of critical thinking, moral reasoning, leadership, well-being, and other dimensions. Post-tests were collected each spring for the survey using the National Survey of Student Engagement (NSSE) and the Wabash Student Experience Survey (WSES). A sample is examined with 5,743 students in 47 institutions, tracking across the fall and spring semester, looking at six dimensions on the Scales of Psychological Well-Being and the impact of participation in community service.

In another study, Padgett et al (2012) use a Wabash sample of 2,609 students to examine the impact of the first year of college on undergraduate students as part of Project Muse. The results suggest that first-generation students "are significantly at a disadvantage in cognitive and psychosocial measures" (p. 252), as well as in their development of a Positive Attitude Toward Literacy. First generation students "may benefit in different ways from exposure to the same empirically vetted good practices in undergraduate education" (p. 261). These students are "underprepared to interact with faculty upon entering college" and "a level of discomfort and intimidation may exist, which can be overwhelming for any student" (p. 261). They should be encouraged in high school to seek help and have academic discussions with their teachers, preparing them to be "less intimidated by interaction with college faculty."

Ewell (2008) describes an example of SUR longitudinal study in the NCHEMS report entitled "Tracking Postsecondary Students across State Lines: Results of a Pilot Multi-State Data Exchange Initiative." Funded by Lumina Foundation for Education, NCHEMS brokered a four-state exchange of data from State Higher Education Executive Officers (SHEEOs) in Kentucky, Ohio, West Virginia, and Tennessee. Ewell explains the benefits of using multiple states' SUR data to study migration. Citing Adelman, he notes that approximately 40% of students transfer across state lines. Unit record data "can be used to track student transfer," but "they have rarely been linked to examine patterns of interstate mobility" (NCHEMS, 2008, p. 1). While the results show only modest migration representing a boost in graduation rates of approximately one percent, retention across states is increased by one-half to three percent. "Bachelor's degree completion rates for border institutions are slightly, but not appreciably, enhanced by including other states. Somewhat larger increases are apparent for retention rates…" (p. 7).

As part of its migration study, NCHEMS examined data from the National Student

Clearinghouse (NSC). While "Data reported to NSC were never intended to be used for cohort tracking…, many institutions report enrollment and degree completion data to the NSC each year" (Ewell, 2008, p. 7). NCHEMS and NSC staff were able to "construct GRS-type cohorts for five states" and generate statistics on retention and completion at the same school, same state, and all institutions. The results show an increase of approximately 7% in graduation rates if all sources of data are included.

Data from the pilot data exchange were expanded by Milam (2008) with SUR data from Virginia and the State University of New York (SUNY) system to create a six-state "Student Unit Record Exchange" (SURE). The SURE dataset includes sixteen fields: record type (enrollment or completions), encrypted student identifier, IPEDS school identifier, enrollment term/year, birth year, gender, race/ethnicity, county of origin, student level, attendance status (full- and part-time), and CIP Code for the current field of study. Completions records include data about the term and year of award, CIP Code, and award level. Awards levels include one and two year certificates, associates, and bachelor's degrees. All enrollment and completions records were tracked for the fall 1998 and fall 1999 cohorts of first-time students for 18 semesters through spring 2005. The two cohort years were combined into one large group and rates calculated for the six-year period. Eight measures are examined – majors, dropouts, transfers, retention, persistence across institutions, and certificate, associate's, and bachelor's degree awards. The degree progress between certificate, associates, and bachelor's degrees was also calculated.

Focusing on nontraditional student success in public two-year colleges, Milam (2008, 2009a, 2009b) finds that the nontraditional dropout rate is 41.9% versus 32.9% for four-year. The first-year transfer rate is lower for nontraditional students (3.9%). "Nontraditional students are more likely to earn a certificate (3.1%) or an associate's degree (18.2%), especially in the two-year cohort (4.2% and 21.5% respectively). They are much less likely to earn a bachelor's degree (9.9%), though this varies from two-year (3.9%) to four-year (22.0%)" (Milam, 2009a, p. 1). "Mixed attendance, a combination of full- and part-time enrollment, dramatically increases nontraditional student retention, from 41.6% for full-time to 76.3% for mixed. Bachelor's attainment is only slightly higher for mixed (10.5%)" (p. 1). "Some nontraditional majors have retention rates that are 20 to 30% above average for this type. Nontraditional student bachelor's rates are higher in some areas of national need, including education (33.5%), math (35.5%), the physical sciences (22.8%), and social work (22.0%)" (p. 2).

"Associate's attainment for nontraditional students drops from 23.2% for those who continuously enroll to 16.2% for two enrollment spells and 8.1% for three or more spells" (Milam, 2009b, p. 2) Also, "The certificate rate for nontraditional students in STEM is 8.0%, compared to 4.2% for all majors. In areas of national need, this rate goes up to 14.9% in agricultural sciences and conservation, 10.8% in the health professions, and 10.6%

in nursing. These are higher than rates for traditional full-time students" (p. 2).

Prescott and Ewell (2009) present a vision for a multi-state "human capital development data system" for sharing SLDS data and highlight the types of questions that may be addressed. These involve "Tracking the stock and flow of the skills and abilities (represented by education and training) of various populations within a given state" and "Examining the gaps in educational attainment between population groups, based on demography and socio-economic status" (p. 3). This system is being put in place with P-12, postsecondary, and workforce data from Hawaii, Idaho, Oregon, and Washington (Prescott, 2012).

Based on the availability of different databases in one or more states, different questions may be asked. In a single state, what proportion of beginning students earn the bachelor's in six years? What proportion complete high school and enroll in college within a year? What proportion of high school graduates complete college within 10 years and earn 35K or more per year? If data are shared across multiple states, it is possible to ask: What proportion of students enrolled in one state enroll in another the next year? What proportion complete high school in one state complete an associate's degree and are employed in a specific industry in the same or other states?

### STATE STUDIES

NCHEMS, SHEEO, DQC, CCA, and others document state progress in building P-20 longitudinal data systems. Every state has interesting examples of using these data to understand student success, evaluate programs, and ensure continuous improvement.

In Illinois, this capability is being implemented by the Illinois Education Research Council (IERC) as it works on various studies across the P-20 range of education (http://www.siue.edu/ierc/). Illinois' focus on the LDS is also reinforced by a state level P-20 Council. The P-20 Council was created by the Illinois legislature in 2009 and has as its mission "to deliberate and make recommendations to the Governor, Illinois General Assembly, and state agencies for developing a seamless and sustainable statewide system of quality education and support, from birth through adulthood, to maximize students' educational attainment, opportunities for success in the workforce, and contributions to their local communities." (http://www2.illinois.gov/gov/P20/Pages/About.aspx#what ).

Several statewide examples help to illustrate the range of questions that may be addressed with a longitudinal database, along with some of the constraints.

Conklin (1995) describes a five-year, longitudinal study of first-time, full-time students at community colleges in Kansas that incorporated eight surveys between 1985 and 1990. The study hoped to determine variables that affect progress, gain insight into attrition, and better understand stop-out behavior that might be masking as attrition. Perceptions, opinions, and experiences are documented and tied to four indices of institutional effectiveness – transfer, persistence, educational goal attainment, and satisfaction with the college experience.

Ewell (2008) describes case studies of using longitudinal data in the North Carolina, Kentucky, and Florida community/technical college systems, as well as two- and four-year systems in California. In partnerships with ETS and ACT, the relationship between placement test results and course grades was analyzed in North Carolina. As a result, placement test cut scores are used more effectively to establish proficiency.

A longitudinal study conducted in Florida found that students who completed a student development/ life skills class in their first year of college had a much greater rate of completion Ewell (2008). These student development (or College 101) courses are now being required in some systems across the country. Now that these have been implemented in many places, questions are being asked about "the timing of students taking the course, and what bearing that might have" on persistence, credits earned, grades, milestones, and graduation (AIR, 2010, p. 52). Questions are also being asked about whether the "worth and promise" of these courses are paying off. In a recent Community College Research Center (CCRC) analysis of student-level data about student development courses at three Virginia Community College System (VCCS) institutions, Karp et al (2012) find that "contextual factors made implementation challenging and undermined the courses' potential to create long-lasting impacts" (n.p.). Sometimes, there are few opportunities for "in-depth exploration and skill-building practice" and these courses need to "include pedagogies that promote applied learning, contextualization, reflection, and deliberate practice."

Longitudinal research on Florida community college students examines the relationship between the rigor of high school classes and developmental needs in college. As a result, some systems around the country are providing feedback to high school districts about the college readiness of their graduates. For example, the Virginia Community College System provides data for each high school in each school division for each jurisdiction. There is also an Illinois High School to College Success Report. The data include the number of graduates; the number of graduates entering each college; the number that were previously dual enrolled; grade distribution by subject; placement rates in developmental math, reading, and writing; pass rates in developmental English and math courses; types of curriculum; and majors/program plans. Data are reported annually for the current year and five year totals. These reports are what Terenzini (1987) calls "autopsy studies" that look backward at behavior and help explore what happened.

In California, the correspondence between the high school curriculum and the expectations of college have been studied using longitudinal data with programs such as the California Partnership for Achieving Student Success (CAL-PASS). Faculty in both settings are working to map the curriculum and improve the transition between high school and college. This makes more effective use of the senior year in college and helps ensure more success. Rodriguez et al (2012) use longitudinal data on dual enrollment courses in the Concurrent Courses Initiative in California to examine pathways from dual enrollment to career and technical education. Hughes et al (2012) conduct a three-year study of California dual enrollment students, focused on under-achieving and under-represented students.

The Community College Research Center at Teachers College, Columbia University (CCRC) has conducted a wide range of research projects involving "quantitative analysis of student success and completion." Numerous questions are addressed, including developmental education, academic preparation, college transitions, distance education, accountability, and other topics. Funded by Lumina as part of ATD, as well as by Gates and other foundations, CCRC is a leader in using state SURs for longitudinal studies. CCRC also produces an "Assessment of Evidence" series that uses "the research literature to draw conclusions and provide concrete evidence-based recommendations to practitioners, policymakers, and researchers" on different topics that benefit from longitudinal study (CCRC, 2012, p. 1).

For example, using longitudinal data from the Washington State Board for Community and Technical Colleges (WSBCTC), Prince and Jenkins (2005) found in a CCRC study that nontraditional students who complete a year of college work and earn a certificate or other credential within five years end up with a substantial boost in income. This can be the difference in helping them climb out of poverty and is what the authors call a "tipping point." The results help motivate policymakers to develop better programs for adults with basic skills making the transition to college. The researchers document the data dictionary of SUR data needed to evaluate other possible tipping points. Ewell (2008) reviews the study and cites it as a noteworthy example of longitudinal work, noting that, unfortunately, "few students reach the tipping point."

Leinbach and Jenkins (2008) track WSBCTC cohorts of first-time and transfer students starting in 2001-02 over multiple years and document the completion of milestones. They identify specific momentum points toward meeting them. A typology of six student groups was developed and used to look for patterns in attaining milestones. The typology includes English as a Second Language (ESL), Adult and Basic Education (ABE), developmental, college level, vocational, and transfer students. The study identifies specific momentum points for each group. The results are disaggregated by demographics, including race/ethnicity, gender, enrollment status, financial aid, and age.

In a similar study of SUR data on community college students in Florida, Calcagno et al (2007) analyze milestones for nontraditional age versus traditional age students over six years. They find that milestones such as earning 20 credits are "significantly less important for older students than for younger ones," suggesting that credits earned "may not be the best way to measure progress" (p. 793). Enrollment in remediation had a less negative impact on likelihood of graduation on adult learners, which "may reflect the varying motivations and goals of older students" (p. 794). Finally, the study showed the "unexpected finding" that controlling for ability using math placement scores, "older students have a higher probability of graduating in any given semester" (p. 796).

Leinbach and Jenkins (2008) provide technical notes about working with longitudinal data. They recommend taking care in the selection of first-time only students and stress the value of including students who start in spring and summer semesters, not just fall, and of both full- and part-time students.

Note that if the IPEDS definition of first-time student is used, colleges should include previously dual enrolled students, but only after high school graduation. Some schools may incorrectly consider these students as continuing and exclude them from first-time cohorts (AIR, 2010). If tracking is done of dual enrolled students who enter college after graduation, the high school graduation year needs to be used to ensure that student major is correctly applied in determining cohorts.

Dual Enrollment (DE) is discussed several times above as a growing policy interest, particularly for community colleges. Win et al (2012) discuss the use of SUR data in the state of Illinois to understand dual credit/dual enrollment and data-driven policy. They address criticism about quality and oversight of DE programs. They find that "some of the research is conflicting or inconclusive on the benefit of dual credit" and that one recent study in Florida "finds no evidence that taking any dual enrollment course improves marginal students' outcomes" (p. 4). In contrast, a recent longitudinal study of Florida found that "dual enrollment has a positive influence on degree completion for low-income students" (Miller, 2012, n.p.).

Using Illinois SUR data, Win et al (2012) ask whether DE is "associated with an increased likelihood of postsecondary degree seeking enrollment" and how these "associations vary across family income levels, gender, race, and other factors" (p. 10); also "to what extent is dual credit associated with decreased time to bachelor's degree completion" and how these "associations vary across specific academic and socio-economic variables" (p. 10). The entire state, public high school, graduating class of 2003 is analyzed, with data brought together from ACT, the Clearinghouse, IPEDS, and Barron's Profiles.

The results of a Cox regression model suggest that DE has "a significantly and strong effect on all forms of postsecondary enrollment" with the effects "somewhat larger for higher income students" (Win et al, 2012, p. 14). The effect, particularly at community colleges, "rivaled that of high school GPA, race, and some of the regional differences" (p. 17). However, "community college enrollment was only significant in terms of predicting an accelerated time to bachelor's degree for students from the low income group" (p. 18).

Another noteworthy example of regional longitudinal study was conducted by the Indiana Project on Academic Success (IPAS), which has spawned a number of regional and institution-level longitudinal projects as part of an initiative funded by Lumina Foundation for Education. IPAS is part of the Wabash studies and is a contributor to several Clearinghouse's Signature Reports about transfer behavior.

The National Center for Postsecondary Research launched the Learning Communities Demonstration project in 2006 to look at the effectiveness of one-semester, learning community programs. The Opening Doors project within this focuses on communities for students in developmental education. Longitudinal data about learning communities at six community college sites in this project, five including developmental students, are discussed by Visher et al (2012). These programs link two or three courses, including one in developmental math or English, and include support services such as extra tutoring. There was variation in the colleges on the developmental subject targeted, the courses that were linked, the integration across the curriculum of the courses, and the use of additional support services.

Using longitudinal data across three semesters, it was found that learning communities "had a small, positive effect on progress in the subject targeted," but not on credits earned in other courses (Visher et al, 2012, p. ES-5). There was a "small positive effect on overall academic progress" in credits earned and the different sites were "fairly homogenous" in this regard. The results suggest, though, that "learning communities had no effect on persistence" (p. ES-7). Cost data associated with learning communities are analyzed and a rubric is presented that highlights basic, midrange, and advanced implementation of the four components of the learning community model.

A longitudinal study over six years of the effect of Open Doors Learning Communities Program at one of the participating institutions, Kingsborough Community College, is presented by Sommo et al (2012). SUR data from the City University of New York (CUNY) and the National Student Clearinghouse are used to collect data on enrollment and degrees. The findings show that more students in the program earned a degree than did non-participants. There was a positive impact on total credits earned. Persistence was improved during the first four years of enrollment, unlike in the larger study. After six years, however, the "estimated impact on persistence is no longer statistically significant" (p. ES-6). Also, it was deemed that the program was cost effective. The "cost per degree earned was lower per program group member than it was per control group member" (p. ES-7).

### Institutional Context

Reynolds et al (2010) use a mixed methods approach to study the impact of a "life calling" course at a religious institution. IPAS data are used to track a cohort of 1,748 first-time, full-time students from fall 2000 through spring 2005. Student background, academic preparation, college experience, and financial aid variables are included. Group interviews are used to identify emerging themes and qualitative methods are used to validate the results with techniques such as peer debriefing. The results suggest that participation in the class has a strong effect on persistence. The qualitative analysis identifies patterns in which students may have benefited, including leading "more intentional lives," coming "to terms with expectations about having a declared major," and access to and use of support structures.

Wabash study data are used by Trosset and Weisler (2010) to understand persistence, looking at the first year experience at Hampshire College and developing predictors for attrition in the class of 2012. ACT maintains the Wabash Study student unit records and the authors linked these data to those on admissions, enrollment, and progress through the curriculum. Two subscales of the Scales of Psychological Well-Being were identified with persisters. A number of "good practices" have been identified through the Wabash Study and these are analyzed in terms of retention. It was found that "retained students were significantly more likely to think that their instructors frequently gave assignments that helped in learning the material" (p. 83). The college's retention committee used Wabash data to document academic and social experiences "reported more positively by retained students" (p. 87). Reviewing the benefits of longitudinal data, Trosset and Weisler state the importance of identifying "correlates of attrition." This leads to better retention practices, based on data from program evaluation, and to improved admissions screening.

Many other examples of institution-level longitudinal research may be cataloged, but there is not sufficient space in this monograph. The most common application of the longitudinal study, especially among public two- and four-year institutions, is with student tracking.

Online, data applications such as those developed by the Virginia Community College System (VCCS) using a SAS data warehouse, allow users to monitor different success rates, momentum points, and milestones for multiple fall cohorts of students by semester for 10 or more years. Data are brought together from enrollment and award records, as well as the Clearinghouse to document the most current and complete figures on continued enrollment, completion, and transfer. These are expressed as counts and percentages and the online reports may be arrayed with combinations of different column and row variables. The reports may be disaggregated for various student and program characteristics, such as degree level, program, attendance status, developmental placement/status, race/ethnicity, Pell recipients, first generation, gender, residency, prior dual enrollment, IPEDS cohort status, and other categorical variables. Detailed, student-level data may be viewed and exported from any cell in the report, tied to the SIS student identifier in a way that allows researchers to merge in other data from surveys and instruments (Milam, 2009c). These data are analyzed in a series of Student Success Snapshots by the VCCS to examine topics such as the success of dual enrolled students and the completion of momentum points and milestones by subgroups.

Longitudinal data are also used in institutional program review in a variety of ways. This can include tracking full- and part-time cohorts of new students in a program over time, with the length varying by the 100, 150, and 200% of completion time. Retention rates are often calculated and Clearinghouse and other sources of state SUR data are brought together to calculate an overall student success measure that includes completion, continued enrollment, and transfer. Employment data for graduates and nongraduates of a program are included, if available, with details about whether students are working in the program field, average salaries before and after completion, and lists of top employers by industry type. Due to the sheer number of programs offered by many institutions, program reviews of this nature are completed over cycles such as every three years. Concentrations and specializations can be lumped together or broken out, depending upon other factors such as full-time faculty, program leads, and data needed for resource allocation models.

Additional institution-level examples are mentioned in Section IX on sources of data.

## Dissertation Research and Other Scholarly Studies

It should be recognized that longitudinal studies have long been used for dissertation research and theoretical scholarship. The reader should not neglect these important, cumulative additions to the knowledge base as examples of longitudinal study. They often use very sophisticated statistical techniques and reflect the rigor of peer review as part of gatekeeper publications in the field, such as *Research in Higher Education*, the *Journal of Higher Education*, the *Review of Higher Education*, the *Journal of College Student Development*, and the *Handbook of Theory and Research in Higher Education*. These theoretical contributions sometimes embody a lifetime of research and scholarship using longitudinal data to understand how college affects students.

In path analysis, event history modeling, and other multivariate methods used to study theoretical perspectives on persistence and completion and the academic environment, the longitudinal study has held a central place. The reader is encouraged to read Pascarella and Terenzini (2005) for a careful review of the types of studies on students which are possible with longitudinal research. Examples include Smart and Feldman (1998) on the use of Holland's theory of accentuation effects to understand academic departments; Ishitani's (2003) assessment of first-generation student attrition; Stupinsky et al (2008) on critical thinking and perceived academic control; DesJardins (2003) for his use of event analysis to understand the role of financial aid; Long and Kurlaender (2008) on community college pathways to the baccalaureate; and Bryant (2011a, 2011b) on the development of ecumenical worldview.

Recent dissertations coming out of Iowa State, University of Michigan, Texas A&M, Harvard, the University of Arizona, the University of Wisconsin at Madison, West Virginia University, Clemson, Arizona State, Penn State, James Madison, UCLA, Western Michigan, the University of Hawaii and others contribute to this knowledge base and illustrate the importance of longitudinal data and evolving multivariate methods to informing the higher education community.

The reader should note that "Longitudinal Studies" is one of the descriptors used by the Education Resources Information Center (http://www.eric.ed.gov/ ) maintained by U.S. Department of Education. When searched with the descriptor "Higher Education," thousands of articles are identified, as do results of searches in databases of dissertations and theses such as ProQuest Dissertations and Theses (PQDT).

In reflecting on How College Affects Students, Pascarella (2006) describes ten directions for future research. There have been tens of thousands of studies using college student samples, of which "only a subset" is "actually concerned with estimating the net or unique impact of the postsecondary experience on students" (p. 508). There are now approximately 2,400 plus studies of college impact being done per decade, representing a "dramatically increasing volume of research" with the potential for ten thousand more studies in the next 10 to 20 years. As this momentum builds, Pascarella recommends these directions:

1. Focus on the quality of the data or information being analyzed – "If we have learned anything from existing evidence on college impact, it is that good data trump almost any other consideration – including the use of sophisticated statistical procedures" (Pascarella, 2006, p. 509)
2. Reassert the importance of replicated findings
3. "Expand our notion of diversity" (p. 511)
4. Acknowledge the increasing diversity of the American postsecondary student population by estimating conditional effects
5. Bring systematic inquiry to bear on the rational myths of higher education – "there is a tendency to base policy decisions on what some have called 'rational myths'" (p. 513)
6. Extend and expand inquiry on previously ignored students and institutions – "there is intriguing evidence to suggest that the academic and out-of-class experiences that influence intellectual and personal development during college differ along such dimensions as race/ethnicity and first generation" (p. 514)
7. Investigate the full range of impacts of information technology
8. Conduct studies that uncover the "why" of an intervention's impact – "the total effect, even when statistically significant, yields little information as to the underlying processes or mechanisms that account for the effect" (p. 515)
9. Map the role of within-college experiences on life after college; except for the "economic effects of college major and grades, almost no attention has been given to mapping the long-term impacts of specific within-college academic and nonacademic experiences during college" (p. 516)
10. Continue to take periodic stock of the research literature to establish where we are and where we need to go (Pascarella, 2006).

## Section 5. Benchmarks and Performance Measures

Multiple measures of student success are needed within the context of multiple missions, explains the Committee on Measures of Student Success (CMSS, 2011). This is well understood and has resulted in the proliferation of a large, sophisticated, and growing research agenda about the improvement of graduation and retention rates. This is being driven by federal and state governments, national associations, and accrediting bodies with an expanding interest on the part of institutions in performance measures and the use of results for the continuous improvement of teaching and learning. The following section discusses using a longitudinal student database to calculate benchmarks, measures, and key performance indicators. If these approaches are implemented at the reader's institution, the appropriate variables will need to be incorporated in his/her longitudinal file/data mart.

The National Governor's Association (NGA) embraced the Lumina Foundation for Education, the White House, and others' national goals for increasing college completion and developed its own approach to metrics for completion, efficiency, and effectiveness. The NGA recognizes the limitations of the IPEDS data collection (Reindl, 2011) and recommends the use of "progress metrics" to look at success in first year courses, earning credits, retention, and course completion. A set of "outcome metrics" are recommended to look at time to degree and credits to degree with longitudinal tracking. The NGA initiated Complete to Compete in 2010 with a list of metrics to track progress and success across states. Complete College America (CCA) is an extension of this initiative (Completion Arch, 2012).

Addressing issues of economic competitiveness, fiscal constraints, and demographic shifts, many states are "using data to achieve equity in educational outcomes" that require following student progress. Longitudinal cohorts are tracked with key milestones, with the results disaggregated by race and ethnicity to help evaluate progress toward equity and completion goals (Prescott and Bensimon, 2011).

Ewell (1987) states the importance of looking at the "big picture" first before establishing student flow models for the institution or system as a whole. Institution-wide calculations should be disaggregated "until distinct behavioral patterns emerge." Once these data are in place, the analyst must "Revisit the model continuously and check the validity of its assumptions." This is because "Longitudinal enrollment models – even if grounded in considerable historical data – are only as good as the assumptions upon which they rest. These assumptions often change as a result of changes in policy or changes in the institution's operating environment" (p. 18).

The following sections look at three broad approaches to documenting student success with performance measures that move beyond the limitations which have been highlighted above about the Student Right to Know Act methodology collected with the IPEDS Graduation Rate Survey (GRS). The first, related to community colleges, is driven by the recognition that the GRS fails to capture the different types of students served in two-year institutions. A unique set of measures has emerged, fueled by Achieving the Dream (ATD) and other initiatives and these are discussed in detail. The focus of the second section is the documentation of transfer. This is a key function for students in both two- and four-year schools. The debate about accurate definitions of transfer and the types and extent of this behavior are highlighted.

Finally, the third section is a discussion of student assessment measures. This is put forward with all the caveats and converging interests that make the topic of outcomes and student learning so dynamic and changing. What are some of the key performance measures that may be developed using longitudinal data to track assessment? Can impact be fully understood and where and how does the focus on "value added" fit in with these concerns for student tracking?

## COMMUNITY COLLEGE STUDENT SUCCESS

While most of the material in this monograph is applicable to all levels of postsecondary education, there are some specific aspects that are uniquely associated with those institutions that award certificates and degrees less than a Bachelor's degree. These schools will be referred to in the following discussion as Community Colleges although it is recognized that some institutions are technical schools and some of the schools who offer Associate degrees and Certificate of less than four years also offer Bachelor's degrees and sometimes graduate degrees.

The Carl D. Perkins Career and Technical Education Act of 2006 (Perkins IV) funds career and technical education (CTE) programs at the secondary and postsecondary level. This program started in 1963 with the Vocational Education Act. The most current 2007 legislation includes requirements for increased accountability and linkage between K-12 and higher education. As part of receiving monies, states and institutions must document performance using new measures designed to assess the effectiveness of the program (GAO, 2009). Postsecondary measures include:

- Technical skill attainment – career and skill proficiencies, including technical assessments with industry standards if available and appropriate
- Credential, certificate, or degree attainment – industry recognized
- Student retention or transfer – in postsecondary education or transfer to four-year program Student placement – military service, apprenticeship, placement or retention in employment
- Nontraditional participation and completion – nontraditional fields employment, "such as women in automotive programs or men in child development" (GAO, 2009, p. 6)

States and systems that implemented this accountability system started collecting new data in 2008-09. The methods for calculating the measures use relatively simple numerators and denominators. For skills attainment, for example, the numerator is the number of CTE students who earn a GPA of 2.5 or better in the reporting year divided by the number of CTE students. U.S. Census data from the Household Survey are used to establish gender and minority percentages of employment by occupation. Underrepresented is consider 25% or less employment in the field. In the calculations used by the Virginia Community College System, these measures are calculated using the following: "the number of minority gender students who were enrolled in a gender under-represented CTE program divided by the total number of students graduating from gender under-represented CTE programs" (VCCS, 2010, p. 3).

As part of ATD and the Ford Foundation's College Bridges to Opportunity project, a number of performance measures for community colleges were developed, along with information about the data elements needed to produce them. Ewell (2008) documents the measures most states and systems can already produce using longitudinal data, then adds a subset of special measures limited in their availability. He diagrams a "Student Enrollment Pathway" that leads from Adult Basic Education/General Education Development (ABE/ GED) and ESL through the bachelor's degree, along with different milestones. Ewell recommends that subpopulations of first-time students be tracked in cohorts over six years, with breakouts by gender, race/ethnicity, age, attendance status, need-based aid status, family income by location, first generation college, and whether the student is a single parent.

The standard measures put forward by Ewell and others from the ATD research include rates in the form of percentages for:

- Completion of any kind of credential
- Completion of associate degrees
- Annual persistence in enrollment from year to year
- Developmental Success I - success in developmental coursework in any area and separately for reading, writing, and math
- Developmental Success II - progress from developmental work to college-level classes
- Transfer from a transferrable degree program to a four-year school (Ewell, 2008)

There are additional measures that require data on course transcripts and unemployment (UI) wage records. These are calculated annually over time in the form of percentages and include:

- Achieve "College Pathway Status" with completion of one semester of college-level work
- College Path Completion (Credentials) tracked after first semester to earning any credential
- College Path Completion (Degrees) tracked after first semester to earning an associate's degree
- College Path Persistence tracked after first semester the students who remain enrolled
- Developmental Success III - from completion of developmental work to completion of first semester of college-level work
- Developmental Success IV – from completion of developmental work to completion of college-level English composition and math courses (separately)

- Noncredit Conversion – from entering the cohort as noncredit students to enrollment in classes leading toward a credential, separately for degree and for credential programs and for those starting in GED, ABE, and ESL
- Basic Employment – from first enrollment in classes leading to a credential to the third quarter of Unemployment Insurance (UI) wage records after completing a credential, separately for those with and without earning a credential
- Post-Enrollment Earnings – Unlike other percentage rates, this is the estimated annual earnings of a student as of the third quarter of employment after earning a credential or after the last known enrollment in a credential program (Ewell, 2008).

ATD initiated a "Cross-State Data Initiative" to develop better ways to measure community college success and a number of intermediate milestones are recommended for tracking (Brown, 2009).

- First-year milestones: persisted fall to spring; passed 80% or more of attempted hours; and earned 24 or more credits.
- Second-year milestones: persisted fall to fall; completed developmental math by year two; and earned 48 or more credits.
- Third-year milestones: passed gatekeeper English or higher and passed gatekeeper math or higher.
- Fourth and sixth-year measures: award of less than associate's degree without transfer; award of associate's degree or higher without transfer; award of less than associate's degree and transferred; award of associate's degree or higher and transferred; transferred without an award; continued enrollment with 30 or more credit hours; and total success rate.

Discussing next steps, Brown (2009) notes that performance should be disaggregated and analyzed according to academic readiness, income, race/ethnicity, and gender. It should be understood that "not all measures pertain to all students." The selection of measures needs to reflect the mission of the state, system, and institution and reflect differences in priorities and policies. For example, Florida encourages transfer after earning a degree, while Texas does the opposite. Some states such as Ohio may not have strong transfer policies, while others such as North Carolina take a balanced approach. At the institution level, longitudinal data are particularly important for identifying at-risk students, improving student advising, and reviewing policies and practices such as course-taking sequences and drop/add policies. "Course sequencing" can be analyzed with longitudinal data to understand course-taking behavior. With this information, researchers may "find courses that are barriers or gateways to student success" (AIR, 2010, p. 40).

Developmental education indicators are addressed by Jobs for the Future (JFF) as part of a larger framework and strategy.

There needs to be "an economical set of indicators" focused on improving the outcomes of developmental education. These should be tied to employment outcomes if possible and disaggregated by age, income, race/ethnicity, gender, and level of developmental education need. Enrollment, completion, and pass rate in development sequences are considered key intermediary measures (JFF, 2009).

Many of these JFF and ATD measures are included in the VFA for community college tracking with longitudinal data. Additional VFA measures include, but are not limited to: percent of credit hours completed in the first term; percent reaching full- and part-time student credit hour thresholds by the end of year two; percent of credit hours completed by the end of year two; percent that left the institution without an award and without transfer that earned 30 or more credits and that earned fewer than 30 credits. Credit hours completed versus attempted is defined as earning a C or better grade or pass grade if the course is pass/fail.

The VFA also highlights Career and Technical Education (CTE) measures for students who complete credit and noncredit programs or that leave college with at least 90 contact hours. These CTE measures include: number of awards; licensure exam pass rate; and percent of students completing a program or 90 contact hours that are employed with a livable wage.

It should be noted that the VFA in the draft discussed here represents an ideal system. The American Association of Community Colleges (AACC) included a variety of colleges in a pilot study about how the data could be collected and the methodology is being revised based on this feedback. Some of the data elements could not be developed and some are so onerous and burdensome as to be cost prohibitive. There is also the question of which measures would actually be used critically by decision-makers.

Jobs for the Future (2007) describes "essential features" for measuring performance. These include: (1) an economic set of measures tied to priorities; (2) longitudinal data to track key benchmarks; (3) measuring levels of college readiness; (4) controlling for student characteristics such as enrollment status, age, and socioeconomic status; (5) controlling for institutional characteristics to permit comparisons; (6) incentives tied to goals to increase success; (7) expanding the timeframe for benchmarks to take into account part-time students; (8) identifying "intermediate predictors" in the first and second year that predict future success; (9) reporting results to identify strong success with "high-priority subgroups"; and (10) incorporating the latest research in adopting benchmarks for predictors.

Noncredit activity measures are included in the VFA as well. These include noncredit workforce course enrollments, number of state/industry-recognized credentials, and the percent of CTE students that transition from noncredit to credit classes. In tracking the percent passing licensing exams, the VFA measure includes two breakouts: graduates/completers who pass the license/exam "on their first attempt" and "within one year of completing the CTE program… separately for each exam" (AACC, 2012, p. 36).

Adult Basic Education/GED measures are included in the VFA for different population cohorts. These include: the percent of students completing the ABE/GED; the percent of ABE/GED students that enroll in further education; and the percent of these students that then gain employment (AACC, 2012).

"The Completion Arch" (n.d.) is a research project and online data application that is hosted by the College Board that brings together national data from different sources. These include IPEDS and the longitudinal BPS study from NCES; state longitudinal data from disseminated studies conducted under SLDS; the Institute for Higher Education and Leadership Policy (IHELP); the Columbia Community College Research Center (CCRC) at Teachers College, Columbia University; the Complete College America (CCA); and other research conducted as part of ATD. Data definitions are incorporated from these sources, along with the VFA from AACC and Complete College America.

With the online completions data tool, the user selects indicators and states of interest, then may download and view spreadsheets containing state- and sector-specific data where they are available. Except for IPEDS, almost all of the data come from longitudinal SUR studies collected from association, foundation, federal, and state projects and submissions.

The Completion Arch includes these longitudinal measures from national and state sources:

- Placement in, participation in, and completion of the first developmental course and the developmental sequence (ATD, VFA, CCRC, CCA, PETS);
- Enrollment in and completion of gatekeeper courses (ATD, PETS);
- Threshold number of credits in specified time (CCA, VFA, PETS);
- Persistence over semesters and years (BPS, CCA, IPEDS, VFA);
- Completion of the transfer curriculum (not identified yet; see NSC Signature Report 4);
- Full-time attendance in the first semester (ATD, IPEDS);
- Completion of courses attempted (PETS, CCA);
- Specified credits within one year (CCA, VFA, PETS);
- Continuous enrollment (ATD, CCA, PETS);
- Summer credits earned (PETS);
- Graduation rates (IPEDS, CCA); and
- Completion rates within six years (IPEDS).

An additional measure mentioned by AIR in its longitudinal training module is program retention. This cohort tracking descriptor includes students who have "taken sufficient courses to be considered as being committed to a program or defined Major/Area of Study" (AIR, 2010, p. 25). This also helps in determining student intent to obtain a degree, certificate, or transfer. "The procedures for performing course retention studies normally mirror those for persistence studies" (p. 25).

## TRANSFERS

The following section discusses some of the research that has been done on student transfer. If the reader is interested in conducting research on student mobility, you will want to collect, code, and create supplemental data with derived variables about where your students go and what they did after they left the institution.

"Definitional problems have bedeviled attrition and retention studies for decades," explains Terenzini (1987, p. 21). The concept of "dropping out" is "really a matter of perspective." Some students may be getting exactly what they want from an institution, then drop out. Their definition of success may not involve completion of a credential or attendance for a full academic year.

Definitions of persister, stop-out, dropout, and attainer were emerging in 1987 when Terenzini published his chapter in AIR's Primer on Institutional Research. Definitions are still evolving 25 years later, with the widespread use of Clearinghouse and state SUR data. Types of withdrawal must be differentiated and analyzed, including internal attrition (transfer between program majors) and institutional attrition (leaving a particular college entirely).

Conklin (1995) describes two methods for calculating transfer rates. The first, which she attributes to the Center for the Study of Community Colleges at UCLA and Art Cohen, tracks the entering student cohort. The second is what she calls the "National Effectiveness Transfer Consortium model with BW Associates" and tracks an existing student cohort.

Romano and Wisniewski (2005) describe the many types of transfer statistics that can be calculated using Clearinghouse data. They cite the Cohen methodology, noting that it includes a subset of first-time students who complete at least 12 credit hours at the community college, and dividing this into the number that take one or more classes at a four-year school within four years. In several key studies, this transfer rate runs "somewhere in the mid-20% range" (p. 2). Adelman (1999) suggests that 30 to 60% of all bachelor's recipients have taken one or more courses at a community college, implying that the transfer process is not linear, but a "swirl" of students moving in and out of institutions.

Transcript requests have become outdated and are no longer a useful indicator of intent, especially at two-year institutions. "Incomplete records of student movements result" (Musoba et al, 2008, p. 102). Only with multi-institutional data such as state SURs and the NSC is it possible to know if, when, and how transfers occur. This means that "we have almost as many definitions of transfer as we have studies of the process" and "the definition used by any particular study is most likely dictated by the data available" (Romano and Wisniewski, 2005, p. 3). Rates should therefore be calculated separately for transfer and non-transfer programs or majors, if possible.

The Consortium for Student Retention Data Exchange (CSRDE) defines transfers as those transferring with 30 or more semester hours from a community college. It then tracks retention and graduation rates for cohorts of these transfers at a 4-

year institution. This combines full-time and part-time students.

Transfer rates at community colleges in the State University of New York (SUNY) system are calculated in the longitudinal study by Romano and Wisniewski (2005). The authors explain that "colleges can track almost all of their own students who transfer into both public and private colleges and across state lines" (p. 2). The availability of Clearinghouse data improves the accuracy and data integrity of these studies considerably.

Using data from the NSC "more than doubled the number of community college transfer students" that could be identified in the Romano and Wisniewski study (2005, p. 9). This suggests that "current research has underestimated the transfer rate by at least 25%. This raises the transfer rate of the Cohen cohort from the mid-20% range to about 30%. Considering the limitations of the NSC data, we might be able to add another 5% to that number" (p. 10).

The difficulties of determining transfer behavior are outlined by Porter (2002), who uses NSC data to analyze transfer-out behavior. "Many students categorized as stop-outs are actually transfer-outs" and these need to be treated differently than true stop-outs who "decide their educational goals are best met by discontinuing their education altogether" (p. 3). Porter describes how state SUR data might incorrectly classify students as stop-outs because they transfer to private institutions where data are not collected. He explains that surveys of graduating students about their "intent to transfer" are incomplete due to low response rates. Monitoring transcript requests is a costly and laborious process that typically is not practical. Only with NSC data is it possible to adequately account for transfer-outs.

Using multinomial and binomial statistical models, Porter analyzes problems in each transfer approach and the ability to predict outcomes. Breaking down attrition into stop-out or transfer-out, he finds that "two variables still have a significant impact" – application time and unmet financial need. "Students who applied late and students with large unmet need are both more likely to either stop-out or transfer" (2002, p. 9). While it may seem counter-intuitive, "First generation students are less likely to stop-out, while in-state residents and participants in an honors program are less likely to transfer" (p. 10).

State SUR data in Indiana were used to report back to institutions on the difference between students who transfer versus those who withdrew. Musoba et al (2008) conducted "inferential analysis of factors that predicted transfer," including GPA and major before and after transfer and attendance status (p. 103).

As mentioned, student departure is a longstanding topic in the institutional research literature. It needs to be acknowledged that "many studies that attempted to faithfully replicate Tinto's model often failed to incorporate the departure timing in their research. Thus, they overlook how dimensions of time affect student departure behavior" (Ishitani, 2008, p. 108).

In one of several recent NSC "Signature Reports," Hossler et al (2012a) study transfer and mobility activity for a fall 2006 cohort of first-time students over five years. This was part of research conducted in conjunction with the Indiana Project on Academic Success (IPAS). The NSC was able to provide a "national near-census of student enrollments" for this cohort. The team at NSC defined transfer as "any enrollment after the end of the student's fall 2006 term in an institution that is different from the institution in which the student was originally enrolled, provided that the student had not already completed a degree or certificate and was not still enrolled (concurrently) in the original institution" (p. 16). Even short-term enrollment at another institution, then returning to the original institution, is categorized as a transfer in this study. Approximately 33.1% of the cohort transferred at least once.

The Signature Report shows that the "transfer-out rate" is higher than national estimates obtained with IPEDS. Part of this is due to the five years' length of time used. The results are also higher because "institutions do not pursue nonreturning students to identify transfer activity" (Hossler et al, 2012a, p. 17). In addition to showing that one-third of all students transfer at some time before getting a credential, "more part-time students transferred than full-time." Of all students who transfer, "37% transfer in their second year; 22% transfer as late as their fourth or fifth years; 25% transfer more than once; 27% transfer across state lines; and 43% transfer to a public two-year college" (p. 5).

Student outcomes by semester are discussed by Lillibridge (2008) in documenting persisters, stop-outs, completers, and transfers. A cohort file is merged to these other datasets and flags (variables with markers) are created in semester data over time so that unknown outcomes can be better documented. New variables are created to store the three potential values of outcome: completing, persisting, or unknown, depending upon updated enrollment activity over time.

Lillibridge likens the visual display of changes in attendance over time to a "slot machine," because "it is easy to imagine a slot machine with a wheel for each semester. The possible student outcomes used by the tracking model are placed in sequence on each wheel. The values of each wheel click into place as the sequence steps are processed." Stop-outs are determined as "Logical if-then statements are again used to establish this value for each student in the cohort" (2008, p. 24). Students who have multiple outcomes in a semester where they both persist and complete require the researcher "to decide the order or sequence in which variable values are established. The most important ones are established last" (p. 25). Certificate and degree completers are marked accordingly with this approach, using a "trumping rule" in which earning a degree is considered more important than earning a certificate.

It is more difficult implement this logic when there is the "swirl effect" identified by Adelman. There are "students who may transfer and return, earn a certificate and a degree, or earn a degree and then return to college to take more classes. The model deals with these conditions by creating additional values. The complexity of the logical if-then statements increases as more outcome variable values are added to the model" (Lillibridge, 2008, p. 27).

Research by Hossler et al (2012b) in another NSC Signature Report, funded by Lumina in conjunction with IPAS, suggest that there are many more reverse transfers than one might expect, with four-year students going to community colleges in large numbers. This explains "what happens to many of their 'missing' students" that four-year colleges don't have a way to track." Almost 1.3 million beginning students at four-year institutions in fall 2005 are tracked.

"Reverse mobility cannot easily be classified as either a positive or negative student outcome" (Hossler et al, 2012b, p. 10). Multiple reasons are documented from the literature: (1) to save money; (2) changing educational goals; (3) summer course work taken closer to home; (4) personal situations requiring a move home or elsewhere; (5) poor academic performance; (6) difference in welcoming environments of four- and two-year schools; and (7) flexible access to multiple institutions in a region.

The results show that 14.4% of first-time students at four-year schools transfer to two-year colleges within six years and an additional 5.4% only take summer courses. Part-time students and those at public institutions are more likely to reverse transfer. Only a small portion of these students (16.6%) returned to their original four-year college and more than half (55.1%) did not return to any four-year college. Regardless of intent, the majority of reverse transfer students do not return. Most stay at the two-year institution for more than one semester. "While conventional retention studies would categorize them as nonpersisters, the result nevertheless shows that these students continued their postsecondary career and earned credentials in the two-year sector" (p. 6). When students take courses at two-year schools just for the summer, their four-year completion rate goes up almost 20% to 77.5%.

Differences in transfer rates at community colleges based on a career-oriented versus liberal arts curriculum are addressed by Deng (2006). Using SUR data from the City University of New York (CUNY) system, a sample of students who graduated over time from one community college and transferred to four-year CUNY schools is tracked. Approximately a third of the graduates were in liberal arts. Senior college GPA is the dependent measure. The transfer rate for liberal arts graduates is 8.8% higher. However, career-oriented graduates "earn a significantly higher GPA" (p. 5). Career-oriented programs "are not designed for transfer purposes and transfer from these programs to a four-year institution requires a bilateral rather than being based on a statewide articulation" (p. 6). This suggests the need to "rethink the design of career-oriented programs" to ensure "a secure and smooth transfer," with the need to possibly address accreditation criteria for faculty credentials and adding general education coursework.

There is often a problem, Pai et al (2008) explain, with the designation of transfer students because they "bounce between two or more institutions, even enrolling in classes at two institutions during the same semester." The use of a student status flag is therefore inadequate. Sometimes, the way transfer students and those "new to program" are coded is done in ad hoc ways, "without documenting or communicating such changes."

The CMSS recommended that IPEDS graduation rates be expanded to include students who achieve within 100%, 150%, and 200% of normal time the following status:

a. Earned an award; transferred to a four-year institution without an award; or were substantially prepared for transfer
b. Earned an award and did not transfer to a four-year institution
c. Earned an award and transferred to a four-year institution
d. Were substantially prepared for transfer
e. Are still enrolled at the institution in the first term immediately following the tracking period or transferred to another two-year institution within the tracking period (CMSS, 2012, p. 20)

At the IPEDS Technical Review Panel (TRP) convened in response to the Committee report, it is explained that students "who transfer to another institution after being awarded a degree or certificate are not counted in this transfer-out rate" in the GRS. While there is "value in collecting more data on outcomes," the TRP "was concerned with data limitations and the availability of data necessary to report such as rate" (IPEDS TRP, 2012, pp. 6-7). State SUR data are only available to public institutions and do not always include private schools. There is a cost associated with accessing Clearinghouse data and some SHEEO agencies do not have NSC data at their disposal.

One of the most interesting recommendations of the CMSS is the inclusion of transfers that are "substantially prepared" by an institution. This addresses the recommendation made over time by JLARC, ATD, AACC, and others to combine enrollment, completion, and transfer rates into a single "student success" rate. Without this combined rate, transfer is "significantly underreported in part because institutions have limited access to the necessary data" (CMSS, 2012, p. 19).

The definition of "substantially prepared" is problematic. There is a "lack of a clear and consistent definition of what is meant" and "whether it fully addresses all transfer preparation activity across all institutions" (IPEDS TRP, 2012, p. 7). Thresholds are recommended for documenting the number of credits that are attained at the originating institution. However, these and other course-taking behavior milestones and momentum points require sophisticated longitudinal work that must be supplemented with each new semester's data. If substantial preparation is to be reported, four-year schools would need to use NSC data to document the number of their students, for whom they awarded a specific threshold of credits, who then transferred to other four-year schools or reverse-transferred to two-year schools.

Clearinghouse data are often as complete as, and certainly help to supplement, state SUR data. State SUR collections do not uniformly include all private institutions. Private institutions in Illinois are required to participate in the SLDS as a requirement of receiving state financial aid monies. State SURs don't always include data on out-of-state institutions such as for-profits that don't operate out of brick and mortar facilities in the state.

Online offerings in particular are subject to some degree of regulatory oversight. In October, 2010, ED "program integrity" regulations were put forward that would have required all institutions to document their online courses delivered to non-residents (WCET, 2012). The regulations would have required that colleges have some kind of memorandum of understanding (MOU) in place with their non-resident students' home SHEEO agency to ensure accountability. SHEEOs put forward sample MOU language, but the regulatory environment for this reporting was halted by ED in summer 2012.

This focus on institutions to measure the extent of online delivery and degree completion by non-residents has to move forward in some fashion, as it is a requirement of receiving Title IV federal financial aid that all institutions be approved to operate in the states where they serve students. The Lumina Foundation funded a project by the Council of State Governments and the Presidents' Forum, a collaboration of institutions service adult students, to draft a State Authorization Reciprocity Agreement (SARA). Without it, the burden on institutions would be very great (CSC, 2012). See a working draft of the agreement at: http://www.csc.org/NCIC/documents/WorkingDraft.pdf. Note that, while this larger discussion is taking place, there are new IPEDS data collection requirements for 2012-13 that include data on enrollment by delivery mode. Eventually, this could include delivery mode by residency, by which ED could ensure that agreements are in place for states where online students are served.

## Tracking Time to Completion

The IPEDS GRS has been expanded to include 100%, 150%, and 200% of expected time to completion in the calculation of rates. Even 200% may not be sufficient and Long and Kurlaender (2008) and others suggest that ten years of data are needed to completely track success; longer, if employment and non-credit and continuing education/workforce training data are to be incorporated. The CMSS recommended that NCES use sample survey data to "determine a time-to-degree period that would capture graduates at appropriate timeframes without imposing too great a burden on institutions in tracking several cohorts over many different timeframes" (CMSS, 2011, p. 16).

In expanding the IPEDS GRS to include a cohort of part-time students, it was acknowledged that 200% of time may not be adequate. "Normal time to completion… is defined for full-time students, but not for part-time" (IPEDS TRP, 2012, p. 8). "Given the variation in part-time students' enrollment intensity and patterns and the undefined normal time to completion for these students, the panel suggested that the revised GR200 component collect data at yearly intervals" (p. 8). Three options for reporting cohorts were considered, requiring up to eight years for four-year institutions and six for two-year.

Cohort tracking is exacerbated for institutions such as career and technical colleges that "offer programs that enroll students on a continuous basis and with programs starting throughout the school year" (p. 12). Note that ED's Gainful Employment (GE) collection got around this problem by collecting an entire year's worth of enrollment and completions data at a time, without regard to the semester label.

National data on the number of terms that students are enrolled to completion are provided in a Snapshot Report on Mobility report by the NSC (2011). However, time to degree is a much more complex discussion than may be gleaned from what has been reviewed so far. This topic has its own "small but growing literature" (Knight, 2004). An interesting example of this research using longitudinal, institution data to inform IR is provided by Knight in an AIR IR Applications paper entitled "Time to Bachelor's Degree Attainment: An Application of Descriptive, Bivariate, and Multiple Regression Techniques." Two dependent measures are used – semesters elapsed and semesters enrolled prior to degree attainment. Demographic, pre-college, enrollment behavior, academic outcomes, financial aid, learning community, first-year program participation, parental education, and program accreditation data were "assembled into a series of data files." Descriptive, bivariate, and multivariate analyses were conducted and the results discussed both as how they inform practice but also as methodological considerations.

There are other ways to count the "time" element towards a degree besides the two Knight uses, it is noted by the IR Applications editor. "The point is that the selection of a dependent measure and how it is measured is not a foregone conclusion. The presence of Web courses and independent paced courses adds to the complication" (Knight, 2004, p. 14). There is also an issue with transfer courses taken outside the institution. "How do you use prior experience as an independent variable or set of variables? Do you need to split the transfers into different sub-groups and look for different models?" (p. 15).

## Student Typologies

The reader is encouraged to think about incorporating typologies to understand variation in student success and explore sub-populations of student groups. Bahr calls this getting a "bird's eye view." Typologies may be created for a variety of purposes, including enrollment behavior, theoretical constructs such as student engagement, and program interventions. "Classificatory activities are fundamentally critical to social sciences" and "are naturally seen as an intrinsic component of knowledge discovery in the data mining process," explains Luan et al (2009, p. 3). There "is a high level of interest in methods of differentiating and identifying types of community college students" (Bahr, 2011, p. 33).

Some student typologies are based on demographic characteristics such as age, gender, geographic location, and/or socio-economic backgrounds. Other typologies are based on responses to various instruments or patterns of activities. Some typologies are based on course-taking patterns and some are based on a mixture of these characteristics.

Milam (2008, 2009a, 2009b) describes a typology informed by Ewell, based on analysis of data for the entire population of first-time students in six states for six years for the SURE project. After extensive exploration, five student types emerged: (1) traditional, college age students who start and continue full-time; (2) traditional, college age students who start part-time, but attend with a mixed enrollment pattern of full- and part-time; (3) nontraditional students, identified as such because they are age 20 or older before starting college and have a mixed attendance pattern; (4) students who attend only part-time, but for more than one semester, regardless of age; and (5) incidental students of all ages who attend college part-time but only for one semester in the six years.

Extensive research on typologies has been done for some time, using data from CIRP, NSSE, and other sources. Hu et al (2011) provide a useful review of this knowledge base and highlights of continuity and change in typologies. Luan et al (2009) discuss "Using a Data Mining Approach to Develop a Student Engagement-Based Institutional Typology" in the AIR IR Applications series. Eight student clusters are documented using nine dimensions that emerged with factor analysis. This typology includes: (1) high-interaction; (2) tradition-learning-focused; (3) homework-emphasized; (4) diverse-and-spread; (5) meeting-service-needs; (6) disengaged; (7) collegiate; and (8) easy-pass.

A behavioral typology of first-time students at 105 California community colleges is presented by Bahr (2010), based on course-taking and enrollment patterns over six years. Cluster analysis suggested a six cluster solution that includes: (1) transfer; (2) vocational; (3) drop-in; (4) noncredit; (5) experimental; and (6) exploratory. There is "considerable variation in the representation of important demographic characteristics within and across clusters, particularly with respect to students' race/ethnicity and age" (p. 741). Bahr notes that "a sizable percentage" of first-time students have "short durations of enrollment." These students "have little exposure to the community college" (p. 742). Where the "experimental cluster attempted more units of coursework but succeeded in fewer courses," the "students in the drop-in cluster appear to be achieving the ends that they seek." Transfer and getting a credential just are not goals for them. These students take fewer courses but "succeeded in those courses at a rate than exceeded the average success rate of students in any other cluster" (p. 742).

## Assessment

Using a longitudinal database for assessment often involves collecting survey data as well as demographic data. The following discusses some of the major work done using various instruments such as the National Survey of Student Engagement (NSSE) and the Cooperative Institutional Research Program Freshman Survey (CIRP). Student responses on these instruments are typically returned to the institution and can be merged with the student longitudinal record. This facilitates using a longitudinal file for research and for decision support.

The use of survey data in longitudinal studies is described by Ewell (1987) in terms of questions to ask at three points in time. At the time of entry, typical questions involve student goals and intended duration or persistence, college choice, and college readiness and the need for remediation. Currently enrolled students are asked about progress toward goals, use and evaluation of programs and services, involvement with the campus environment, perceived gains in skills and knowledge, and perceived changes in attitudes and beliefs. Former students are asked about goal fulfillment, subsequent enrollment, employment, use and evaluation of programs and services, perceived strengths and weaknesses of the educational experience, and reasons for attrition and persistence. These "questionnaire surveys pay their greatest dividends when designed in concert with record-based student tracking studies" (p. 17). Enrollment data are needed to supplement the loss of survey data, due to low response rates. This helps alleviate the problem that survey data are "notoriously unreliable in forecasting or accounting for enrollment behavior" (p. 17).

Endo (1992) recommends that IR practitioners "construct the educational experience questionnaire before the freshman questionnaire… The single best predictor of any student outcome after one year is the value of the same outcome at the start of college. The wording of outcomes measures should be identical on both questionnaires" (p. 30).

The Academic Experiences Study at the University of Delaware is described by Bauer (2004) and is an example of how an institution can develop its own instrument. The survey is used to "examine intra-individual change in cognitive skills, changes in college activities from freshman to senior year, and the relationship between change in personality typology and change in critical thinking" (p. 79).

A number of instruments have been developed that help explain the longitudinal nature of the undergraduate experience. These are based on the cumulative knowledge base of work done by individuals such as Alexander Astin, Bob Pace, Vincent Tinto, Art Cohen, John Bean, George Kuh, Ernest Pascarella, Pat Terenzini, Ken Feldman, John Smart, Corinna Ethington, Vince Tinto, John Bean, Stephen DesJardins, Sylvia Hurtado, Jeff Milem, and others to understand larger theoretical frameworks such as academic integration and student engagement.

The Higher Education Research Institute's CIRP Freshman Survey and College Senior Survey are among the most prominent approaches to this type of assessment (Padgett et al, 2010). These and other instruments from ETS, ACT, NCHEMS, and others have been used for years as part of longitudinal studies of impact (Terenzini, 1987). See the AIR *Handbook of Institutional Research* chapter on "Institutional Research with Published Instruments" by Noble and Sawyer (2012) for a review of the many types of instruments that are available and how they may be used.

The Cooperative Institutional Research Program (CIRP) freshman survey was established by Astin and has been conducted since 1966. Keup (2004) explains that the survey's "most important contribution to assessment in higher education is its ability to serve as a pretest for subsequent longitudinal follow-up of entering students" (p. 8). The follow up instrument can "capture the development of the whole student by covering a variety of areas, including academic achievement, skills, and engagement; learning strategies and pedagogical practices; residential and employment experiences; interactions with peers, faculty, and staff; campus involvement; satisfaction with curricular and co-curricular activities; patterns of behavior; student values and goals; self-confidence and feelings of personal success; and plans to enroll for the second year" (pp. 8-9). These two instruments have been used in "countless studies" with national, consortium, system, and institutional comparisons.

As an example, see the work of Hurtado et al (2008) using longitudinal CIRP and first year college survey data. The authors look at first-year, minority students and key predictors in their participation in structured opportunities for health science research. It is recognized that "These experiences are particularly important for Black students" (p. 126) and the training of future scientists. The results inform programs to "orient students at an early stage" for careers in biomedical and behavioral science research.

In their presentation about "Linked Longitudinal Data Sets," Wittstruck et al (2002) describe Missouri's use of its ACT Data Base with data on approximately forty thousand high school graduates who take the ACT each year. The database contains test results, demographics, high school coursework, and family background data.

Several instruments have been developed to help measure progress over time according to student engagement theory, among them the Beginning College Survey of Student Engagement (BCSSE) and the National Survey of Student Engagement (NSSE). The Survey of Entering New Student Engagement (SENSE) and the Community College Survey of Student Engagement (CCSSE) were developed for use with specific populations and institutional mission. A faculty version of the instrument is also available. Cole and Korkmaz (2010) describe using the BCSSE and NSSE with a sample of fifteen hundred students from institutions around the country. The relationship between engagement in high school and college is explored.

The merits of these instruments to truly measure impact is not the subject of this monograph. With the NSSE, programs to create learning communities and promote service learning are tied to engagement and the impact of these initiatives can be evaluated using longitudinal data. The use of NSSE in institutional research is the subject of an NDIR volume edited by Gonyea and Kuh (2009). Other work being conducted by the American Association of Colleges and Universities and others focuses on high impact practices and how they contribute to teaching and learning.

Nevertheless, it should be recognized that there is "very little internally valid evidence with respect to the predictive validity of the NSSE. This is a serious concern if participating postsecondary institutions are asked to consider the NSSE benchmark scales as a proxy for student growth in important areas" (Pascarella et al, 2010, p. 2). The authors use data from the seven liberal arts assessments conducted for Wabash study and map them to the measures in the NSSE benchmark scales. The NSSE benchmark scores "had a significant overall positive association with the seven liberal arts outcomes at the end of the first year of college, independent of differences across the 19 institutions" (p. 4). At least one benchmark had a significant association with each of the liberal arts outcomes except for the "Need for Cognition" scale. Only the Student/Faculty Interaction benchmark "failed to have a significant partial correlation with at least one of the seven liberal arts outcomes" (p. 5).

Pascarella et al (2010) further explain that the NSSE results are "good proxy measures for growth in important educational outcome such as critical thinking, moral reasoning, intercultural effectiveness, personal well-being, and a positive orientation toward literacy" (p. 5).

A number of studies include grade data. Herzog notes that "Pascarella and Terenzini (2005) eschew course grades as object measures of learning," but that Astin and other still use grades in national sample studies because they are "strongly correlated with standardized test scores after accounting for academic subject, grading variation, teacher ratings, and certain student behaviors" (Herzog, 2011, p. 22). After reviewing this literature, Herzog writes that "course grades coupled with test scores can be used to capture cognitive growth in students that reflects academic achievement over time as well as postgraduate skill level" (p. 23). An interesting example of using grades is the longitudinal analysis of semester GPAs after transfer to four-year schools by Ishitani (2008). Ishitani asks "How do transfers survive after 'transfer shock'?" Sophomore and junior transfers are less likely to drop out than freshmen. As expected, "higher semester GPAs were positively associated with higher persistence" (p. 403). Pike et al (2011) examine first semester grades of students in learning communities.

# Part II: Construction, Design, Display, and Data Sources

Longitudinal data require time and effort to create and maintain. The expenditure of these resources makes sense only if the result adds value to the institution that is committing the resources. The first part of this monograph focused on the use of longitudinal data – why they are important and the kinds of questions they can be used to answer. Part 1 also provided a growing awareness of the additional types of data that might be gleaned from institutional records and/or gathered with supplemental surveys and instruments.

The second part of this monograph focuses on more technical aspects of a student longitudinal database. It is intended to identify key technology and methodological issues. It also provides an overview of the many technical aspects that must be considered. It is not be possible, because of space constraints, to go into all of the practical, how-to detail needed in many cases; but that detail may be found in the references. Also, some of the material will be technical and it is assumed the reader has a working knowledge of database management and the software and hardware used. Throughout the monograph, a number of acronyms are used and these are included in the Glossary that is part of the Appendix.[4]

The basic construction of a longitudinal database is the topic of the first section of Part 2. The discussion moves to obtaining and organizing data and then to typical data structures for pulling all of the different sources of information together to meet needs over time. Issues in the manipulation of the data once they are stored, such as linking tables in a relational model and creating different dimensions and measures, are the next topic. Data integrity concerns are focused on, including issues such as missing data, dirty data, and working with derived and calculated variables. The critical need for a unique student identifier is discussed, along with ways to define and track cohorts of students. The definition of cohort characteristics such as degree-seeking can be muddied and this is the topic of another section.

The use of crosswalks and taxonomies for categorical variables or dimensions such as CIP Codes for disciplines of majors, is addressed. Methodological concerns in designing the longitudinal database must be understood, especially if any kind of advanced statistics are to be used in order to make empirically sound generalizations. Particular attention is given to the construction of local surveys and instruments, to problems in the potential calculation of "value-added," and to program evaluation, sampling, and the use of triangulation and multiple methods. The display of data through technology, including expectations for dashboards and data visualization, are discussed, including the use of new business intelligence tools and open source alternatives for this software. In evaluating these IT planning issues,

[4] An excellent reference on data management is the DAMA guide to the data management body of knowledge, Mark Mosley and Michael Brackett (eds.), Technics Publications, Bradley Beach NJ, 2010.

observations are put forward about the change process that is involved and how to get stakeholders' support.

The many different, potential sources of data from across the continuum, from P-12 to workforce, are described, along with examples of their longitudinal use for current hot topics of interest. Finally, planning issues in developing, maintaining, and staffing longitudinal studies are put forward for consideration. The Appendix includes a comprehensive Glossary of acronyms and references for further study.

## SECTION 6. HOW TO BUILD LONGITUDINAL DATASETS

There are a number of practical considerations in building a dataset for longitudinal studies that may be used to calculate all of these different rates and measures. Ewell (1987), Endo (1992), Lillibridge (2008), and others describe these from an institutional perspective. The ideal state SUR data system at the state level is explored by Pai et al (2008), Ewell and L'Orange (2009), and others; as well as by documents from the Data Quality Campaign (DQC) and the implementation of the Creating Opportunities to Meaningfully Promote Excellence in Technology Education and Science (COMPETES) Act of 2007. Larger questions about longitudinal research in education are explored by Singer and Willett (2003) and others. Concrete aspects of constructing data files are described in software programming guides, such as *Longitudinal Data and SAS* by Cody (2001), and user group presentations such as Priest and Collinsworth's (2011) "Intro to Longitudinal Data: A Grad Student 'How-To Paper." Professional development workshops and presentations at the AIR Forum, such as those by Guerin (2009) and Jones-White et al (2008) about constructing longitudinal data and using them for IR, are also helpful.

This monograph discusses some of the key issues to think about if the reader is building a longitudinal database. Many foundational works cover this material in detailed manner and the reader is encouraged to build a personal library of resources. Four contributions stand out, however, to help explain theoretical, institution, and state perspectives on this process.

Ewell summarizes core questions to address before building a longitudinal database: "Who will be tracked?" "How long will they be tracked?" "How often will new cohorts be established?" and "What data elements will be tracked each term?" (Ewell, 1987, pp. 9-10).

"Retention Tracking Using Institutional Data" is described by Lillibridge (2008), who discusses the types of studies that might be done with a longitudinal data system. In addition to calculating student flow, persistence, and completion, he uses a tracking model to produce IPEDS Graduation Rate Survey data. Using demographic data for the larger cohorts for retention and completion reported on in the Graduation Rate Survey (GRS) many sub- populations may be defined and these provide an important starting point for various studies. Of course, the GRS is not designed to "delve into questions related to course-taking patterns, movement through remedial to college-level coursework, course completion ratios, grade point averages, or other measures of academic success." By bringing in data on course-taking behavior, it is possible to obtain much more "rich findings possible through transcript analysis" (p. 19). Lillibridge defines the steps for building a tracking model. These include: (1) establish the cohort; (2) determine student outcomes for each semester; (3) conduct the analysis; (4) identify completers; and (5) display results.

In their New Directions in Institutional Research chapter "Developing a Statewide Student Tracking Tool," Pai et al (2008) present a state coordinating board's perspective and practical how-to guide. A Student Tracking System Master Table is constructed to store the cohorts and data elements. The table includes one record per student per institution with records for all semesters at an institution. Since the database software Microsoft SQL Server 2005 is used, there is a limitation in the number of columns to 256, requiring "difficult choices about which data elements would or would not be stored in the master table" (p. 13). Other data elements can be included through joins and views, such as financial aid, fall cohort definitions, and tuition and fees charges. The 256 column limit also restricts the database in the number of semesters for which students may be tracked, so a limit of eighteen semesters is used.

Overall, in constructing a longitudinal database, questions need to be addressed about how the data will be extracted, transformed, and loaded (ETL) from different sources and silos into some form of data structure. Users need to know how the data will be stored, the selection and format of data elements, and how transaction-level data will be transformed into dimensions and cubes. They need to know how and when the data will be updated and the process of merging, coding, flagging, and creating new value-added variables of interest in new or existing table structures. Sometimes, programming can take care of this process, perhaps with visual SQL tools. The desired displays and report structures, with their intended levels of aggregation and drill-down capability, must be determined beforehand if the data structures are to be sufficient for their use at the needed levels of granularity. These require that a data structure be sketched out ahead of time, something which Common Educational Data Standards (CEDS) provides as a model.

Issues in data integrity, the use of identifiers across the continuum of data sources, the definition and use of cohorts, and the use of taxonomies for categorical variables must all be considered; as well as how the time element of semesters, census files, and the differentiated stop and start dates of semester sessions and course offerings will need to be handled.

## OBTAINING AND ORGANIZING DATA

There are a number of ways to obtain data for longitudinal studies. If a data mart or warehouse is available, then clean, reliable data may be pulled with the expectation that they have already been converted from the transaction level to the dimensions and data elements useful for reporting and online display. While the warehouse may not have longitudinal reports available, the existence of historical datasets that adhere to standards for variables and data dictionary formats is invaluable. If historical census files exist and these contain consistent variables and value labels of interest, this is excellent preparation.

Ewell's (1987) chapter in *A Primer on Institutional Research* about longitudinal tracking discusses longitudinal file construction. The greatest problem is "the manner in which student record data is generally stored and accessed in computerized registration files" (p. 7). Student history files available for transcripts and other purposes contain historical information on demographics and course-taking, but only for the most currently available data. Frozen term/semester files are the other source of institutional data, but these contain data that may not have been updated or corrected over time. These term files "are freestanding and are difficult to link together without special programming" (p. 8). A set of "distinct longitudinal files" is recommended "for ease of access and data manipulation, and for maintaining important derived data" (p. 8). A minimum record layout for variables most critical for analysis is provided, much as done for state SUR systems in his later work (Ewell et al., 2003; Ewell and Boeke, 2007; Ewell and L'Orange, 2009).

"Gaining access to institutional data is the first step," with needs to extract data from information systems and then store them securely (Lillibridge, 2008, p. 20). "Virtually all institutions have these data, although the ease of accessing them varies. Each record in a file contains one or more keys that permit joining records from a number of files by matching keys. A key may be a student identification code, a course reference, or another identifier" (p. 21).

Data files may be stored and manipulated in a variety of formats, including SAS, SPSS, STATA, Access, Excel, SQL Server, Oracle, PostGres, R, HLM, Mplus, and MySQL[5]. The selection of database software to store data makes less difference, as long as tables may be merged and aggregated into dimensions and cubes and new variables and flags created. There are several standard approaches or data structures for storing the data – either as combined, longitudinal tables or separately in a series of sequential tables that are then joined to produce views for reports (King et al, 2012).

As explained by AIR (2010, p. 7), "There are many ways to organize datasets, and an infinite number of variables to include… When variables required for a specific study are housed in more than one dataset, each dataset must contain at least one common data element to permit linking records together." Regardless of processor, storage, and memory in the computer used, some "datasets are simply too large to be efficiently manipulated" (p. 14). If the dataset is unwieldy, the researcher may "create smaller datasets for longitudinal tracking studies that contain only the variables germane to the study. As studies evolve, however, additional variables often become of interest, and if they are not included in the original dataset, you'll need to go back

[5] Many of the various packages have open database connectivity (OBDC) allowing basic use of different packages.

and create additional datasets from source files" (p. 14).

## Typical Data Structures

Typical data structures include a header file, which is usually a student demographic file of some kind with one record per student per semester, academic year, or annual year. These files may include:

- a course file with one record per student per course section per semester;
- a course schedule file with one record per course section per semester, including the type of instruction and location;
- a completions file with one record per student per credential per award date;
- a teaching load file with the instructor of record;
- other files as desired such as financial aid, testing, surveys, advising, room utilization, tutoring

These files are frequently stored as tables with multiple rows for the individual records. These tables need to be merged prior to the analysis. Critical to this merging of tables is the creation of a cohort definition file that contains identifiers and cohort characteristics. The focus here is on SUR data. But if personnel such as faculty are to be tracked longitudinally for workload, research, grants, or other productivity measures, then a different type of identifier and set of cohort characteristics will be needed; but the process is essentially the same. See for example the longitudinal analysis of research productivity by new faculty by Perry et al (2000).

Four types of data elements for longitudinal tracking are described by Ewell (1987): (1) fixed that never vary, such as demographics and high school performance; (2) variable, recorded each term, such as credit hours attempted and GPA; (3) semi-fixed that do not vary except occasionally, such as marital status, residency, and major; and (4) derived, calculated for analysis, such as age, cohort membership, and course completion rate. There are variables of interest "that cannot be directly measured, or for which data are unknown," in which case "proxy variables indirectly measure an attribute or characteristics" (AIR, 2010, p. 29).

An example of a derived variable that is used by the Virginia Community College System (VCCS) is the student characteristic "under-represented." This is a variable data flag that a student has one or more of four characteristics: (1) from an economically disadvantaged location; (2) minority; (3) low-income (using Pell eligibility); and (4) first-generation college, itself a derived variables based on data about parental educational attainment. These codes come from admissions tables and financial aid. If desired, the data may be kept with separate variables and flags or as a single flag. Pell eligibility may change over time, such as with attendance pattern, so a decision rule is needed about whether, how, and when this flag will be updated.

If course-taking behavior is to be incorporated into the LDS, then course attributes such as location, room, delivery mode, type of instructor (if instructor data are not linked), course credits, grades, and start/stop dates are useful to include. Some coursework, such as for gifted and talented high school students held on college campuses, have nine-month classes. These are input into a student information system (SIS) with appropriate start and stop dates, but must by necessity in the system be tied to one term[6] (usually fall). The data are not available until after the spring term, however, on final enrollment and grades. For this reason, the user must approach course schedule data warily for programs that do not fit normal expectations. Many institutions offer courses of varying lengths, sometimes called sessions. Questions are raised such as how the success of a five-week course in developmental math compares to that of a fifteen-week version. Are the students better prepared to move ahead to the first college-level math class or better retained the first year? The same data issue occurs with short-term programs for athletes to earn several credits in order to maintain eligibility, a recent focus of accreditation and federal concern.

Census files need to be extracted several times a semester. This is usually done on the first day of class, on an official census date, at the end of the drop/add date, at the end of the semester, and after final grades are posted – however these are implemented. Grades may change after a final extract is taken, so a decision needs to be made about whether these files will be updated after the data are supposedly final (Lillibridge, 2008). This occurs because "Students may sit in on classes but not register until later in the term or even after the semester ends," requiring care in the extraction, revision, and use of census files (p. 20). See Borden et al (1996) for a useful discussion about setting a census date to optimize enrollment and retention studies.

An institutional model for studying student retention at the University of Hartford is described by Glover and Wilcox (1992) with a database structure that allows for longitudinal study. The first database contains tables of data about admissions, demographics, characteristics associated with retention such as credit hour attainment and GPA, quality measures such as high school rank and test scores, and enrollment behavior over time. A second database records sequential semester data. These two are combined into a third aggregate table structure, retaining the cohort variables of interest. For this study, only students taking 12 credits or more over time are included in the cohorts. Six years of data are compiled. The database manipulation is explained in some detail and several outcome measures are incorporated, including continued enrollment, completion at different levels, and twelve or more credits earned. A menu-driven, executive information system is described. Though built in the early 1990s, this general approach is still current as it has moved to web display, allowing users to select variables and value labels and investigate their impact on retention and completion rates.

In bringing together data from multiple sources for a longitudinal database that will serve many different needs over time, there is a natural inclination to add as

[6] This presupposes that the institution has some type of term (e.g. semester or quarter). If not, some type of calendar coding system needs to be used to identify points in time.

much data as possible. If techniques such as cluster analysis are going to be used to understand factors, "the results of cluster analysis are highly sensitive to variable selection," explains Bahr et al (2011, p. 69). "Researchers must be wary of using every variable that is available in a given data set simply because it is available and instead focus on selecting variables that are pertinent to the research questions of interest… For example, if a researcher is seeking to understand students' enrollment or course-taking patterns, demographic variables are not dimensions on which the clusters should be projected. Instead, one might consider course credit load, number of enrolled semesters, course success rate, number of courses attempted in math or English, and the like" (p. 69).

## Data Manipulation

It is important to prepare for the complex types of data manipulation that are necessary to creating a sustainable longitudinal data model. In constructing what they call a "Master Table," Pai et al (2008) discuss the step-by-step process of merging data elements into the table from different sources in the data warehouse. While this is at the SHEEO level, the issues raised and general approach are the same.

The first step is "to centralize and consolidate existing student unit record data tables" from the warehouse into the tracking database (Pai et al, 2008). A separate table is used for each reporting year. A Stored Procedure in Microsoft SQL Server is used to run predefined SQL statements. SQL Server Data Transformation Services (DTS) are then used to execute the Stored Procedures.

In one of the DTS packages, a sequence of tasks consolidates demographic and enrollment data elements and loads them into the master database. The reader is referred to this NDIR chapter for more detailed information, but this brief description shows how careful and sequential the construction process must be. The steps include: (1) load credit data with demographics, starting institution, school type, credit hours, and semester GPA for a year and semester; (2) load remedial flag on taking remedial, ESL, math, English, and reading; (3) load number of courses taken by level; (4) load number of course withdrawals; (5) update student status for first-time students who started in the summer to count them in the fall; and (6) update student status for those students who are continuing from term to term.

The second DTS package typically has the following steps: (1) delete cohort by reporting year; (2) insert new cohort for year with designation of students as first-time, transfer, readmitted, or new to program, along with demographic and enrollment data; (3) update flags for students who change cohorts; (4) reset current year tracking, summer, and remedial data when student information changes from when they were first loaded; (5) reset current year completion information for the most recent and accurate; (6) update existing cohort information; (7) update summer cohort information; (8) update cohort degree information; and (9) update status flags for those students who are put into different cohorts or are enrolled for longer than 18 semesters (Pai et al, 2008).

There are many data manipulation techniques which must be done in compiling and using longitudinal data. Where the SHEEO is concerned with collecting consistently defined data from multiple institutions, the institution-level staff person is concerned with extracting consistent data over time with variable definitions and formats that hopefully do not change. The need for referencing and updating metadata[7] cannot be overstated. With metadata and data element dictionaries at hand, the work is ready to begin. Not until the data are merged, recoded, aggregated, and appended to other structures do some of the practical issues come into consideration.

If the data are being maintained in a relational data structure with tables, the user may spend significant time developing a series of merges and table structures based on dimensions and measures of interest or envisioning "functional tables" that bring all of the most interesting data together in one flat, spreadsheet-type form. A certain amount of exploration is necessary to understand the impact of different types of joins and the need to "loop" through multi-record formats to create an artificial single-record. This may be due to the way the data are stored and coded over time in the system or due to multiple transactions such as a series of enrollments registered by an individual prior to a semester. Even when clean, final, end-of-term, official, census data are used, the process of merging datasets, creating new derived variables, and exploring how the data may be aggregated, grouped, and displayed will test the assumptions and patience of the developer. It is helpful to have discussions ahead of time with some of the anticipated users of reports before the data structures are finalized.

Conversations with potential users and discussions of likely uses will help shape many of the design questions that need to be addressed. For example, what will happen when the student's major changes over time? Will you pick the major from when the student was first enrolled or the last enrollment? Will this be overridden by the major/program associated with the award of a completion? What if there are multiple completions in a year? What if the programs or majors offered have changed, been recoded, been collapsed, or expanded? Which effective change date and value label will be used? Will the table include only the highest award or all of them over time? If so, based on what criteria (such as sorting alphabetically or a hierarchy of levels that requires that a lookup table be created and used in another merge)? Many of these answers will depend on how use of the data is envisioned.

When data are linked to employment records, SSNs must be used; but what happens when there are multiple SSNs for a system-wide student identifier? Which should be used? One option is to remove students who have multiple social security numbers. If the student record is removed, isn't that a form of imputation? If there are multiple employers in the same quarter of the same year, which do you pick? Do you aggregate the records so that there is one primary employer and if so, what do you do with records for the

[7] Metadata are data about data. Typically these will include the source of the data, the data type, allowable values, transformations on the data, and where the data are archived.

employer that have multiple street addresses all with slightly different spellings? It is often time and cost prohibitive to clean up these data. Will salary outliers be included that could potentially skew the results? If outliers will be removed, on what basis and how will this methodology be consistently applied over the years of data when the Consumer Price Index (CPI) may have changed? Do you take records where wages are above the minimum wage or poverty level? Or do you sum the wages by year for one report on annual wages by graduates and aggregate them by North American Industry Classification System (NAICS) two-digit codes for another report by industry type of program graduates? Do you report just on students who started in a cohort at a particular point in time? If so, do you use the program majors at that point in time to document program success in workforce measures?

These are just examples of a few of the dozens of design questions that must be addressed with an iterative process as one creates longitudinal datasets and begins to use them for analysis. When the user has experience with these processes over time, some wisdom develops through lessons learned and reflection on the amount of time that was spent and what was accomplished. It is at this point that new taxonomies and lookup tables are created for consistently identifying and using new categorical variables. Whether the end result is a single functional table that appears to users like a spreadsheet or a complex data structure with dozens of tables and relationships, the issues are still the same. Some of these questions are:

1. What variables will be selected and what are the keys that document the table relationships for how tables may be merged?
2. How will the data be selected or filtered for different analyses using dimensions of categorical values or thresholds and calculations of continuous variables or measures?
3. How will the data be grouped for display or aggregated into a new data structure?
4. Is the desired grouping or cube data structure efficient for the reporting and displays that will be used?
5. How will the cube be refreshed or replenished with new data and categorical values?
6. What new categorical variables values are useful for analysis and reporting and how will these be stored, refreshed, maintained, and updated?
7. How will parameters be passed to the grouped or aggregated/cubed data to select and drill down/up on different levels of aggregation using the desired groupings?
8. How might the data be presented to novice users as dimensions and measures?
9. How might the data be presented visually so as to accommodate different selections of variables, filters, and groups in a way that can provide the ability to view additional variables and drill down to different levels of detail?

## Data Integrity

When data are brought together, there can be major problems in "adapting data definitions, data collections, methods of reporting, and timing of reports" that the researcher needs to "systematically document" (Pai et al, 2008, p. 8). The standard approach for understanding changes in data elements is to maintain a data element dictionary that includes all types of metadata. The CEDS project is an example of the specificity that may be used.

In addition to documenting continuous variables or dimensions and the type and range of values that are stored, the values and meaning of categorical variables must be kept with a time/date stamp or effective date that allows for these values to be updated and maintained over time as part of version control. This will help the user determine which values are most appropriate to use for which files and time periods. Metadata tools enlarge the concept of data dictionaries with contextual information. Without these, the process of keeping track of longitudinal data is frustrating and endless. One way to keep track of metadata is with wikis, a sort of online, participatory encyclopedia. Wikis allow a community of users to work together documenting and evolving standards, definitions, and formats (Thomas, 2004b).

Problems may exist with missing data for key variables of interest from student records, such as race/ethnicity, high school rank, high school GPA, and test scores. Musoba et al (2008) talk about using secondary sources to find these data. They note that "While challenges posed by missing data and data management are important to weigh in considering use of SUR systems, a host of ethical concerns must also be addressed" (p. 112).

One such problem can occur in SUR systems when private institution data are collected. The universe of institutions will change accordingly and "in order to avoid having incomplete, and therefore useless, cohorts in the tracking system," a "filter" must be used to remove institutions from the view or join that are no longer in the collection (Pai, 2008, p. 9). This issue occurs due to the changing universe of institutions with mergers, closures, and acquisitions of new colleges and new campuses which may or may not have federal Title IV Program Participation Agreements and state approval in place to offer financial aid.

Another problem occurs when the names and formats of variables change across waves of longitudinal data collection. It is helpful to "have a unique identifier for each type of variable to help researchers decipher what will be relevant for their particular research questions. For example, variables found in a data set that consists of data taken at four different time points are labeled according to the corresponding year" (King et al, 2012, p. 46).

"Nonuniform definitions are a barrier to the creation of any multi-source data system" and have "an impact on the structural integrity of the data tracking system itself," write Pai et al (2008, p. 10). Unfortunately, the "loss of institutional history" due to staff changes can make it "impossible to determine for certain whether or how a data definition has changed" (p. 11).

One approach to missing data that is "simpler yet statistically indefensible" is simply deleting records that have one or more variables with missing data (King et al, 2012, p. 52). While it might seem intuitive to select the cases for which complete records exist, this will strongly affect the results and the user is warned against it. Instead, flags may be put in place so that types of outliers may be omitted from analyses.

Missing data of interest may include race/ethnicity, particularly when the collection of these data moved to the OMB-mandated two-question format, breaking the possibility for trends over time except at a rudimentary level. It is important not to contaminate the coding of this key variable with attempts to reconcile changing codes in a way that cannot be undone. It is always important to be able to go several steps backward in the data manipulation process to reconstruct previous work.

For national NCES studies, imputation techniques such as "hot deck" are used, where a randomly selected, similar record is used to create reliable data that are critical to national estimates. With SUR data, "The data that are present can be used to predict what the missing values will take. The most common method is the expectation-maximization (EM) Algorithm" (King et al, 2012, p. 52).

In collecting student SUR data from institutions, it must be recognized that "data definitions [are] inconsistently applied by reporting institutions" (Pai et al, 2008, p. 9). "The more hands the data pass through before being used in a particular application, the more potential there is for implicit errors" (p. 10). Over time, as institutions use "their own interpretations of how to account for changing student behavior patterns," alterations and expansions of data element definitions can occur (p. 10).

Numerous issues contribute to problems in data integrity and "researchers should be prepared to accept some dirty data," explain Pai et al (2008). "Even if all known definitional variances can be accounted for," there are still programming errors, typos, and other sources of errors that, while small, "do not fall into a pattern" (p. 11). "Sometimes it is best to accept that no data set is ever clean and simply move on" (p. 11). Edit checks may be run as part of the data collection and these should be continuously upgraded to address new problems that come up. Cross-checks between data elements and analyses of discrepancies are standard ways to cleanse the data as part of ETL. It must be understood that "previously overlooked problems can become apparent when data are used in a new way" (p. 12). John Porter discusses the SUNY System data warehouse and enforcing various business rules to correct data errors before collection submissions are locked. "Business Intelligence and analytics are used by system staff to review the reasonableness and completeness of data before they are accepted for migration into the central data repository" (Milam et al, 2012, p. 449).

In documenting the flow of data, problems may exist, for example, with the flow of developmental education placement information. Initial placements in developmental math and English courses may be kept in the student information system, but "will updated placements be entered when a student moves to the next level developmental or college course? Will the updated placements override the initial placements or be defined as second placements" (AIR, 2010, p. 14).

A similar problem exists when students are allowed to have multiple majors or academic plans. One major is designated the primary, but a second or third value can be recorded in the SIS. Faculty sense of workload may include students who are listed as advisees but are not actually enrolled, or students who have secondary majors but do not get counted in productivity reports because only first majors are included. The major of enrollment may differ greatly between the time of admission, the final term of enrollment, and the program completion/award. Care must be taken when reporting on programs.

For employment analyses such as average salary earned before and after graduation, if a student has had more than one major, the appropriate rule needs to be developed as to the major that is used to categorize the student. What program will be used for the report? The first program plan or major selected at the time of admission? If so, this often changes. If the last plan of enrollment is desired, then the programmer must loop through all subsequent enrollment records to select the last value. If it is only the program of award, where a student in an associate's degree earns a certification in general education, then a decision about intent must be made. If only the last major is used, then only the last major needs to be in the database. For other rules, all majors may need to be in the database with start and end times. These are some of the decision rules regarding changes in key categorical variables that must be made. Some student information systems allow students to have multiple plans or majors, especially if they are seeking a joint degree, so the primary and secondary codes need to be determined.

A single variable may be used that "identifies students' placements in composition" that is "based on results on a standardized test, ACT scores, high school courses and grades, credits transferred from other colleges or universities, and credit for prior learning… Remembering all the component variables that feed into the summary placement variable, and the rules regarding score levels and dates that determine whether or not values can still be used to determine placement… can be daunting. When staff turnover occurs, the memory of how the derived variable was calculated can more easily be lost. So it is essential to document the definition for each variable and record how derived variables were formed in a data dictionary" (AIR, 2010, p. 13).

Outliers in the data are discussed by AIR in its longitudinal training module. "Sometimes you won't know about the existence of outliers until data analysis begins. Only then can you make a decision about whether to keep them in the study, exclude them, or include them in a larger group that has similar characteristics" (2010, pp. 29-30). In using employment data, quarterly wages are usually prorated to an annual salary based on the number of quarters and the standard 2,080 work week hours. If quarterly wages are below $2,000 then the figure is below the minimum wage and probably represents a partial quarter that should not be used in annual salary estimates. The sum of wages earned across four quarters is a more reliable estimate, though it cannot be broken down into an hourly estimate. Wages below and above thresholds are suspect for some analyses and need to be addressed. For its workforce success measures, the State Council of Higher Education for Virginia adopted a full-time wage earnings (FTWE) rate of $13,195/year, which is the equivalent of $7.25 an hour times 35 hours per week for 52 weeks (SCHEV, 2012).

AIR (2010) explains how the same variable may be kept in both numeric and character/alpha format in different datasets. This may be due to some inefficiency in the importing process or to a change in the collection over time. Extensive formatting of data is often necessary "to identify mismatches and correct them" and "these technical problems can be challenging" and "make longitudinal tracking studies take longer than it seems they should" (p. 15).

Sometimes, key variables such as grades are dynamic and can change. "Grades change as students make up incompletes or successfully challenge original grades, students are placed back into classes from which they were erroneously dropped, and a host of other reasons" (AIR, 2010, p. 14). These updates come from extracts from transaction systems and be linked back to the original data files to keep the cohort consistent.

## STUDENT IDENTIFIERS

A student identifier is needed to track individuals across different programs, institutions, time periods, and sources of data. Ewell and L'Orange (2009) explain that the Social Security Number (SSN) is "used by most states and can function effectively as a unique identifier," as long as it is "protected, encrypted, and that an alternative identifier is eventually developed" (p. 2). The IPEDS Student Unit Record Feasibility Study recommended an identifier created from student characteristics (Cunningham and Milam, 2005). This national identifier was subsequently portrayed as a kind of student barcode. The National Student Clearinghouse relies heavily on matching combinations of key variables such as name, date of birth, and last institution of attendance. SSNs are required to match data between student records and other key data sources, such as Unemployment Insurance Wage (UI) records. Any use in a longitudinal model needs to be carefully thought through, using SSN when necessary for matching purposes and incorporating a non-SSN identifier whenever possible.

Access, security, privacy, confidentiality, and FERPA protections related to the collection, storage, use, and sharing of SSN and other identifiers are not the purview of this monograph. The reader is referred to the following resources:

- Book Three of the National Forum for Education Statistics' Guide to Longitudinal Data Systems
- NCES State Longitudinal Data Systems (SLDS) materials at http://nces.ed.gov/programs/slds/
- SLDS Technical Brief "Basic Concepts and Definitions for Privacy and Confidentiality in Student Education Records (NCES 2011-601)

- SLDS Technical Brief "Data Stewardship: Managing Personally Identifiable Information in Electronic Student Education Records" (NCES 2011-602)
- ED's Privacy Technical Assistance Center at: http://www3.ed.gov/policy/gen/guid/ptac/index.html
- DQC's resources on FERPA compliance at: http://www.dataqualitycampaign.org/resources/topics/13/
- State Data Systems and Privacy Concerns publication from Jobs for the Future, at: http://www.jff.org/sites/default/files/StateDataSystems.pdf

It should be recognized that student ID numbers may change or need correction. Since this is the "key variable used to join or link" records, a table of these changes needs to be maintained and used as part of merges (Lillibridge, 2008). This may occur with international students who are assigned a temporary ID, with students who do not report an accurate SSN, or with data entry or other processing errors.

In particular, the use of K-12 student identifiers is an issue that should be thought through if there is the possibility of integrating the institution's data with these data. As explained by Ewell and L'Orange (2009, p. 2), "The ability to develop interoperability between a state's K-12 and postsecondary data systems is critical to addressing the educational needs that transcend both systems." Unfortunately, and despite the requirement for states to build P-12 SLDS under No Child Left Behind (NCLB), there are both political and technical reasons for problems in linking the two sources of education data (Phillips, 2009). "Most states today do not have data systems that enable this two-way communication; in many cases, there are two separate data systems and they rarely can exchange information" (Ewell and L'Orange, 2009, p. 6). The National Forum guides to LDS address this issue to some degree, but do not conquer the "cultural differences and turf battles (e.g. who 'owns' the data)" (p. 6).

Other identifiers to consider and plan for include course numbers that are unique to the semester/term, course numbers that part of a master course list, course titles, course subject and catalog number combinations, instructor numbers, facility building and room numbers, and section coding schemes. All of these can serve as keys for matching data. Course data may then be used to track substantive change with program offerings by location, accreditation requirements for faculty credentials, assessment by delivery mode, and other academic and administrative tasks[8].

## Cohort Definitions

A cohort of students meeting a common set of conditions is defined by using various data, such as demographics, with different filters such as prior attendance. Lillibridge (2008) describes using the cohort of first-time, full-time, degree-seeking students that are reported under the IPEDS Graduation Rate Survey and adding filters to look at subgroups; for example "Hispanic full-time degree-

[8] Including data from a) home schooling, b) private schools and c) out-of-state schools.

seeking students or first-time students who took and passed a specific course or courses, such as developmental math" (p. 23).

Membership in a cohort is documented as a variable in the longitudinal, student database. Cohorts are typically created for the fall semester to meet IPEDS reporting requirements. Lillibridge (2008) and others recommend including spring entrants too in cohorts. Whatever approach is taken, there needs to be consistent coding of the summer semester and terms as either leading or trailing the academic year to create a combined, annual file. With many collections, the summer leads. However, the use of different session lengths and differentiated stop and start dates to allow customized course schedules makes the use of semester/term labels seem antiquated. The point is to ensure that all activity is captured consistently and that data from different silos, transaction systems, or sources may be brought together and used for derived/value added measures.

Four tracking cohorts are described by Pai et al (2008), based on the student status of first-time, transfer, readmit, and new to program. If a student changes cohorts, new flags are applied and "a new record is generated in the master table to follow the student's educational progress in that new cohort; however, the new cohort information is also captured in the record of the last cohort in which he or she was designated" (p. 16), allowing the researchers to track students changing cohorts. Reported cohorts are then defined as part of an online Student Tracking Tool for transfer-in students, transfer-out students, and high school dual-enrollment.

Other types of cohorts may be defined for different purposes. For example, Pai et al (2008) describe cohorts for students who are part-time, nontraditional age, and have initial majors in STEM fields. The CMSS report states that "The student cohort used in calculating federal graduation rates excludes many students who typically enroll at a two-year institution" (CMSS, 2011, p. 6). The cohort in the IPEDS GRS has "no information on the academic preparedness of students." Cohort data on college readiness and the need for developmental education "provide important context for interpreting federal graduation rates" (p. 6).

Hossler et al (2012a) have used data from the NSC to study transfer and mobility nationally. The authors recognize that "Researchers face considerable complexity in operationalizing the category of first-time student in analyses" (p. 15). For a cohort of new, first-time-in-college students, they were able to "ensure that a student did not show any postsecondary enrollment record, at any institution covered by NSC data, in the four years prior to the student's fall 2006 enrollment," further excluding students who received any credential prior to the cohort start date. One problem, though, is that "NSC data do not include universal designations for class year" and the data may contain students who earned 30 or more credits in Advancement Placement, International Baccalaureate, or dual enrollment programs. The federal definition used for the GRS does not require that the cohort be freshmen, only that they be first-time in college, full-time, and degree/certificate-seeking. Students who earn credits through AP and dual enrollment classes could enter college as sophomores or even juniors, having earned an associate's degree while still in high school.

Also, "because of inconsistencies in the historical depth of NSC degree database records, it is possible that a small number of graduate students are also included" (p. 15). In SUR data, this may occur because graduates from private colleges are not included in the collection and they subsequently re-enroll in public, undergraduate or graduate programs.

The "Completion Arch[9]" report states that "There is considerable variation, however, in how institutions, states and researchers define degree-seeking" (Completion Arch, 2012, p. 6). "Many community college students who might nominally be considered degree-seeking are in fact unsure of their goals" and "more than half of community college students change their educational objectives over time." "Such inconsistencies can limit the comparability of indicators across colleges and states, but several national efforts are under way to address this issue" (p 6). Furthermore, goals change over time.

The CMSS and ED Action Plan both address the need to clarify the definition of degree-seeking. Information gleaned from student intent, course-taking behavior, financial aid, and enrollment by major are mentioned for consideration. The IPEDS Technical Review Panel (TRP) conducted in February 2012 helped NCES understand issues in expanding the GRS.

Part of the discussion focused on the definition of degree-seeking, noting that the definition is "subject to some interpretation." "High school students also enrolled in postsecondary courses for high school credit are not considered degree/certificate-seeking" (IPEDS TRP, 2012, p. 4).

Student applications for financial aid are viewed as "indicating their intent to pursue a degree or certificate," both for federal and state aid. The application for a U.S. student visa required of foreign students is also viewed as intent to pursue a degree or certificate (IPEDS TRP, 2012).

The CMSS suggested that the number of credits students attempted could be used as a proxy for intent regarding earning a degree/certificate. However, the TRP "was hesitant to include language about course-taking patterns in the definition" because it could be challenging to document and represent a burden for institutions. Three criteria are put forward by the TRP to standardize the definition of degree/certificate-seeking: (1) receiving any type of federal financial aid; (2) receiving any state or local financial aid that specifies eligibility as a degree/certificate-seeking student; and (3) obtaining a student visa to study in the U.S.

The expansion of cohorts to include both full- and part-time students is called for; but, just as the definition of degree-seeking is somewhat unclear and ill-defined, the same can be said for attendance status. Some typologies identify a cohort of first-time, predominantly full-time students that includes those who start full-time, but attend over time with a mix of full- and

[9] Completion Arch is a web-based tool developed by College Board to measure the success of Community College students.

part-time enrollment. The category for those who attend only part-time seems more clear, but this needs to exclude students who are incidental or transient students, taking one or two classes once and never enrolling again. The categorization scheme needs to be sensitive to those with a mix of attendance status. This requires a loop through all of a student's enrollment records to determine if she/he is predominantly full- or part-time. Similar work must be done to determine if a student is predominantly on- or off-campus, tied to a particular campus, day or evening, and virtual/online. Some of these use admissions records and do not change except when registration forms change while others are based on course-taking behavior. Decision rules must be defined and consistently applied for when certain labels "trump"[10] others in the logic.

## Crosswalks and Taxonomies for Categorical Variables

The "development of taxonomies for higher education is a much more imprecise process than is suggested" in the literature, explains Milam (2006, p. 52). In using national datasets, "researchers are reluctant to compromise on the level of detail they desire" (p. 55), sometimes leading to frustration when data from different sources and silos are brought together with a grand vision than cannot be realized.

Classification of Instructional Program (CIP) Codes are a primary example of changing taxonomies that are used for different purposes. A number of different disciplinary mappings from six-digit CIP codes may be created. Some of the mappings developed by NCES and by ED's Office of Vocational and Adult Education (OVAE) for national studies and publications include the following:

- two- and four-digit titles associated with CIP codes
- science, technology, engineering, and mathematics (STEM)
- science, mathematics, and research for transformation (SMART)
- career and technical education (CTE)
- critical shortage areas
- academic/career groupings
- nontraditional fields by gender (Perkins IV)
- career clusters and pathways

These taxonomies help with the study of topics such as the persistence and completion of women and minorities in science and engineering and movement through career clusters and pathways. One way to start understanding institution-level data is to extract a list of current and historical major or program plan codes. These may then be "mapped" or "cross-walked" with a lookup table or format statements in programming from major to department. There are many ways to group or cluster disciplines of majors and awards in ways that are useful to longitudinal study.

CIP codes are not always the lowest level of granularity for majors and programs. These codes are used widely as part of

[10] "Trump" applies when a student has been in multiple categories of an attribute, there is a sequence of which one is used. "Best Address" is an example when someone has multiple addresses.

resource allocation and peer comparison for productivity measures and there is a political nature to them that should be considered. Multiple majors or program plans may fall under the same CIP code that is approved or permitted by the SHEEO office. This duplication is intended, even though it makes things more difficult for reporting. If this is the case, then some CIPs can't be used for meaningful internal reporting; but they may be mapped to other program/major codes. Some examples of using discipline data from different, longitudinal data sources are the analyses of salary needs based on course activity, space requirements, and equipment needs in heavy and light equipment-use labs over time. Research universities and others may use the results to argue for more state funding or for different internal resource allocation schemes.

For historical records, Higher Education General Information System (HEGIS), CIP 1990, and CIP 2000 disciplinary codes may be used. As disciplines change, taxonomies must be maintained and revised and this impacts the use of these in longitudinal studies. Problems in the change of CIP Codes from 1990 to 2000 are discussed by Pai et al (2008). Crosswalks are created to move between the old and new taxonomy, but "no crosswalk is perfect, and potential users should be cognizant of that fact" (p. 9). The NCES CIP 2010 Wizard is an online tool that is part of the CIP 2010 website. It lets users compare the CIP codes used in their last three years of completions data to the new taxonomy.

Since CIP 2010 codes are now required for enrollment and completions data collected in IPEDS, this tool is no longer needed for collection. However, it still provides a relatively easy way for researchers to understand the changing nature of the CIP taxonomy and to dynamically explore alternatives for discipline codes that have changed or been removed over time in an historical, longitudinal dataset. Comparable evolution has taken place in the discipline taxonomy used by NSF for the Survey of Earned Doctorates (SED), the Graduate Student Survey (GSS), and other collections. Recent communicate suggests that the proprietary discipline codes in the GSS have been changed to reflect the 2010 CIP Codes. These disciplinary taxonomies must be understood in the context of institutional mission and program complexity.

Definitions used in taxonomies for categorical variables get revised over time and these changes must be addressed. Some recent national examples are CIP Codes, IPEDS doctoral and professional degree award levels, federal Office of Management and Budget (OMB) mandated race/ethnicity codes, North American Industry Classification System (NCAIS) codes, and Standard Occupational Category (SOC) codes.

To use historical data that cross periods when categorical value labels have changed, the user must "create a crosswalk that will render definitions consistent across time and with other institutions. However, this is a particularly time-intensive effort because each instance will require its own solution. Optimally these individual solutions would be tracked in another table" (Pai et al, 2008, p. 11).

NCAIS codes include twenty different industrial sectors that are identified with the first two digits of this six digit coding scheme. From the third to the sixth digit, the user is provided with a more granular depiction of industries. CIP Codes have been mapped to NAICS codes and provide "mechanisms researchers can use to track the transition from major into industry of employment" (Schenk, 2011, p. 3).

SOC codes are maintained by the federal Bureau of Labor Statistics (BLS) and used to "classify workers into occupational categories." The 2010 version breaks out 23 major groups, 97 minor groups, and 461 broad occupations. "Detailed occupations in the SOC with similar job duties, and in some cases skills, education, and/or training, are grouped together" (BLS, 2012). Job titles associated with the detailed occupations are provided in a Direct Match Title File. The next edition of the SOC will be in 2018, timed to begin with other BLS and census surveys.

A CIP to SOC crosswalk was developed by NCES and BLS and analyzes "the relationship of education and training programs to the labor market." Information about the tasks and characteristics of occupations is available in the SOC and when linked to CIP helps with curriculum planning. The CIP to SOC mapping "must indicate a 'direct' relationship, that is, programs in the CIP category are preparation directly for entry into and performance in jobs in the SOC category" (NCES, 2011, p. 3). Some occupations are not mapped to CIPs when postsecondary education is not required. Some CIPs are not mapped to SOCs when the "program is not career related, or because an insufficient number of institutions offer the program, to justify having a CIP code." Crosswalk relationships are "one-to-one, many-to-one, or many-to-many" and "it is likely that one CIP code will map to multiple SOC codes" (p. 3). The CIP 2010 website available at http://nces.ed.gov/ipeds/cipcode/ provides the crosswalk and documentation.

The National Crosswalk Service Center (NCSC) is a federal program funded through the Iowa Department of Education that provides data files and documentation for these and other taxonomies. See http://xwalkcenter.org for more information. Various file formats containing the crosswalk between SOC and NAICS are provided. One report/dataset called Occupational Employment Statistics is available for different dates. This includes the "estimated employment for each industry-occupation combination is included to help determine the size of the occupation within the industry" (NCSC, 2012). It is noted that "The NCSC recognizes the impossibility of developing crosswalk files that are suitable for every use, so the files developed by the center are presented as *prototype* sets of relationships" and users need to determine the "appropriateness of the contents for their particular application." The bottom line is that there is currently no easily implemented crosswalk between CIP and NAICS. Any conversion requires careful decision rules about the levels of aggregation that are possible.

Two additional categorization schemes, career clusters and career pathways, were developed using the SOC crosswalk and are used in career and technical education (CTE) reporting, such as for Perkins IV funding. Career clusters combine occupations for areas in which workers "share similar interests and strengths" and there are currently 16 clusters.

Discipline codes are not just tied to majors and programs. Canada documents the CIP Code for every course at its public institutions and Adelman and others have used CIP Codes as part of mapping transcripts for NCES national studies of course-taking behavior. If available, CIP Codes for courses provide another lens in the use of longitudinal data to understand student behavior and success. These codes are certainly more accessible than using syllabi to understand the sequence of how competencies are developed and to map the curriculum to student learning outcomes.

Longitudinal studies may incorporate cost data, in which case taxonomies for department structures need to be documented and maintained, tied to other higher education data such as faculty workload, budgets, financial aid, room inventory, space utilization, equipment, and revenues. These additional elements may or may not be tied to program evaluation. While not typically leveraged for longitudinal studies, these other sources of non-student data are readily available in the information infrastructure of most institutions through financial resource management systems and human resource management systems.

The reader is referred to the Delta Project for longitudinal, institution-level data related to costs and finance using IPEDS and other data sources. The Delta data are now incorporated into NCES and IPEDS reports and are available for download and analysis with different tools. See: http://nces.ed.gov/ipeds. State and institution-level capacity to conduct longitudinal cost studies with SUR data should not be underestimated (Milam et al, 2008). Longitudinal SUR have been at the heart of simulation and resource allocation models since the 1970s (Milam and Brinkman, 2012). They are integral to the development of an induced course load matrix (ICLM[11]) to display the constraints facing departments and majors.

Similarly, new categories of course activity should not be forgotten, including delivery mode and other types of program offerings such as non-credit, dual enrollment, Adult Basic Education (ABE), English as a Second Language (ESL), Perkins, etc. These may have the same use of student identifiers such as SSN, which is helpful in tracking movement across the continuum. Developmental Education (DE) courses may not be tied to a specific room or location the same way on-campus class activities are. In this case, the data available on course attributes may not be sufficient and it becomes necessary to map other ways to document location. One way is with the use of section codes that document the high school where a DE course is actually taught. Delivery modes and the use of technology are emerging and these are being incorporated into section codes too. Synchronous and

[11] An induced course load matrix (ICLM) arrays data about the consumption of courses by majors against the contribution of courses by departments. It is based on SUR data that combine the level and major of the student with the level and department of the course. Data on enrolled students are then used to "induce" a work load requirement on the department.

asynchronous labels may have been adequate at one time to describe electronic delivery, but these do not describe the full range of delivery modes, especially with Learning Management System (LMS) software such as BlackBoard Collaborate that offer every imaginable type of tool for communication in an online class. Codes for hybrid, online, compressed video, MOOC's and in-person classroom instruction need to evolve and should be somehow be documented as a taxonomy in course attribute data.

Some interesting questions are being asked about student behavior and possible typologies for online learning and how delivery mode impacts success in completion and employment outcomes. Accrediting agencies want institutions to assess whether the experiences of online students are comparable to those served in the traditional classroom. It is necessary to determine the percentage of course work that a student or graduate takes online. Adding up credits earned by delivery mode over time is not as much an issue as documenting program requirements and the range of electives which are allowed. This essentially requires an electronic degree audit, merging major/program plan requirements and course data longitudinally.

## Section 7. Methodological Concerns

While the practical issues of getting and using longitudinal data have been addressed, significant emphasis should be placed on how these data will be used. Many things are possible. From a research standpoint, however, great caution needs to be taken before expecting too much from a Longitudinal Data System (LDS) in terms of policy analysis and research. The complex, multivariate nature of the data precludes their use for generalization, unless the research design permits. With data that encompass the entire population of students over time, almost anything being studied can appear to be significant. The apparently simple analysis of a program's success is fraught with perils for misinterpretation.

This section describes some of these design issues that must be understood. It is not a substitute, however, for a basic grounding in quantitative methods and statistics. This is particularly true when home-grown instruments and surveys are being used, the topic of the second section. The greatest interest and the greatest concern is with the determination of change, impact, or "value-added" and this is discussed in section three. Further discussion follows with sections about program evaluation, restrictions due to sampling, and the use of multiple methods.

### Design Issues

"Longitudinal tracking tells the story of what happens to students, but not *why* it happens" unless the research design allows (AIR, 2010, p. 6). Many studies "can show that two or more variables are associated with one another, but not that one variable causes the other." "Correlation among variables does not equal causation" (p. 12). To establish a "relationship between two or more variables, three criteria are essential: the variables in question must co-vary, the relationship must not be attributable to other variables, and the supposed cause must precede or be simultaneous with the effect of time" (Bauer, 2004, p. 82).

Bauer (2004) explains that longitudinal studies give researchers "the ability to identify individual variations in growth or to establish causal relationships between variables. Collection of data on individuals at three or more points enables powerful statistical modeling techniques, and the precision with which parameters of growth can be estimated improves with each additional wave of data" (p. 79). However, "it is important to understand the distinction between age effects, period effects, and time effects" (p. 81). The researcher "must also consider that contemporary events may have substantially affected individual attitudes, leading to more conservative answers over time. These conservative attitudes are not simply a function of age but the result of societal and intrapersonal events encountered over time" (p. 82). Examples include post-9/11 and the advent of texting and social networking that may particularly impact experiences of the Millennial Generation.

Terenzini explains that "From the outset, it is important to keep in mind that research design is a series of compromises. Designs that increase the power of a study in one area come almost invariably at the expense of some other aspect of the study" (2010, p. 38).

There are a number of flaws in the cross-sectional design used frequently in assessment of outcomes. While longitudinal designs are recommended as an alternative, "Ideally, one would follow over the same period of time a control group of high school graduates who do not attend college… and could be compared after some period of time with the freshman group who have presumably benefited from college attendance" (Terenzini, 2010, p. 39). Cross-sectional studies "invariably reflect characteristics of the same students when they first entered college," explain Astin and Lee (2003, p. 657).

One problem with the use of multiple measures to understand student change is they are "subject to an important threat to internal validity" – "regression to the mean" (Rocconni and Ethington, 2009, p. 369). This is "the tendency to score closer to the mean the second time a test is administered. However, all individuals do not do this. Individuals with highest pretest and lowest pretest scores will tend to be closer to the mean on posttest while those in the middle or 'average' on pretest will tend to randomly shift slightly about the mean on posttest" (p. 369). This "obscures the fact that the error component of the measurements has changed" (p. 369).

With longitudinal studies involving pretest-posttest designs, "special attention should be placed on measurement issues so as not to stray away from internal validity. While regression to the mean may seem problematic, a simple adjustment to the initial score will not only help control for the regression phenomenon but will help protect internal validity and, as indicated above, could result in drastically different conclusions about change" (Rocconni and Ethington, 2009, p. 372). This adjustment involves calculating the correlation between the initial measures and the change between the initial measures and at the end of the time period under study. Adjustments are then made in the initial scores to account for the effects of regression to the mean. The data are then re-analyzed "utilizing the adjusted initial scores" (p. 372).

The study of learning communities as part of the Opening Doors project discussed by Visher et al (2012) is an example of establishing causality by random assignment. It involved random assignment of students to the treatment and the control group at six community college sites. A total of "6,974 students were randomly assigned, about half whom had the chance to enroll in one of the 174 learning communities across the six colleges" (pp. ES4-ES5). Students attended an "intake session" and completed a consent form agreeing to participate. A Baseline Information Form with questions about background characteristics was completed and students were then assigned randomly to the groups. Approximately 71% of students assigned to the program group were enrolled at the drop/add date of the first semester. Different semester cohorts were sampled and combined in the analysis. Student characteristics such as age and gender were similar across the control and program groups.

It is noted that "random assignment does not typically allow researchers to separate the effects of one program component from another" (Visher et al, 2012, p. 27). Separate qualitative research was conducted to provide information on program components. However, this "cannot answer the question of which components mattered most for student outcomes such as credit accumulation and persistence to the next semester" (p. 27).

## Additional Advanced Statistics

Structural equation modeling is useful in representing complex phenomena such as departure, while "doing justice to the impact weights involved," explains Boughan (2000, p. 9). It "treats the relationships among the components of a complex system as a series of multiple regressions overlapping in their independent and dependent variables, the interactions of which are set by the analyst" (p. 9). Boughan applies structural equations modeling and cluster analysis to community college longitudinal data to understand the role of academic process in student achievement. The findings of path analysis and cluster analysis are presented as 10 varieties of academic, student careers, including four achiever clusters, three moderately at-risk clusters, a profoundly at-risk cluster, and two "outrider" clusters. Without this understanding of the academic process, retention researchers can "run the risk of oversimplifying their representation of outcomes phenomena, perhaps crucially compromising the realism of their models" (p. 13).

The timing of interventions to prevent attrition is discussed by Ishitani (2008), who uses structural equation modeling to study longitudinal student departure. While it is interesting to understand how robust different retention theories are for understanding different types of students and institutions, there is more focus on "programs that are effective in improving the first-year retention rate" (p. 109). "What makes their practice more efficient is to know which students are more likely to leave and in what time period" (p. 109).

Event history modeling is used by Ishitani (2008) to understand at-risk students such as first generation and different types of departure. He explains that "Advanced statistical techniques such as event history modeling, designed to assess the process of certain events, can be applied to existing educational issues and topics that have never been investigated in a longitudinal framework" (p. 116). A simulation model may be built that allows the user to select combinations of student characteristics of interest and determine the probability of departure and the expected time to graduation.

Luan et al (2012), Luan and Zhao (2006), Luan et al (2009), and others show how data mining is an effective methodology for analyzing complex longitudinal data for research such as the development of student typologies. The term data mining is "a misnomer. It should be – and indeed was, for a while – called 'knowledge discovery in databases (KDD)" (Luan et al, 2012, p. 478). "In higher education, once the data have accumulated for more than three to five semesters, the database is a mile deep and a mile wide. Without scalable tools and multiple models of data mining, it is impossible to fully understand and exploit all the patterns, trends, and factors in the database" (p. 478).

There are two types of data mining – supervised and unsupervised. Supervised data mining is "when the researcher has prior knowledge of the patterns in the database and is looking for predictive outcomes," using different models to find the optimal predictive outcome (Luan et al, 2012, p. 479). "Unsupervised data mining aims to classify and identify potentially meaningful patterns in data without a preconceived notion of an outcome (dependent) variable" (Luan et al, 2009, p. 3). "To a large degree, this is to uncover the 'natural' existence of clusters in the data" using techniques such as principal component analysis and clustering algorithms. Common software used in institutional research for data mining includes SPSS and SAS.

Chen and Cragg (2012) and O'Connell and Reed (2012), discuss the use of multilevel modeling (HLM) in institutional research using complex, hierarchical data structures such as longitudinal tracking. This approach is "uniquely suited" for longitudinal studies "of students nested within classes or courses, classes nested within departments or schools, faculty within departments, athletes within sport designations within departments or schools – each of these settings describes lower-level individuals (that is, students or faculty) nested or clustered within one or more higher-level contexts or groups (that is, within classes or within departments). In such cases, the variability in lower-level outcomes (student retention, faculty satisfaction) might be due in part to differences among higher-level groups or contexts (class size, department size, and so on)" (O'Connell and Reed, 2012, pp. 5-6). HLM helps researchers "model these dependencies" in order to study variation. It must be recognized that "Observations of student achievement, faculty productivity, and other important performance indicators may be affected by group-level similarities based on these organizational structures" (p. 10).

Multilevel modeling is "still underutilized in the field of institutional research," explain Chen and Cragg (2012, p. 96). The authors provide an example of using HLM to understand the impact of learning communities. They look at how high school performance affects participation in learning communities; whether participation affects persistence, and whether "the types and characteristics of the learning communities have a differential effect on first-year students' GPAs" (p. 97).

Bowman and Small (2010) use HLM to study college student spiritual development longitudinally and Lietz and Mathews use it to understand the "effects of college students' personal values on changes in learning approaches" (2010, p. 65).

Longitudinal datasets with many measurements over time have been labeled "Intensive Longitudinal Data" (ILD) and require additional methodological consideration. The Penn State Methodology Center has conducted research on ILD and what is called a "time-varying effect model" (TVEM). This model "lets researchers see changes in relationships between variables without making assumptions about the nature of those relationships" (Penn State, 2012, n.p.). Examples for understanding ILD are provided by the center's website, such as the dynamics of quitting smoking. "Attempts to quit smoking are influenced by a broad range of factors, including mood, belief in one's ability to quit, and stress level. Using TVEM, we can model the changes in these relationships. This allows us to determine when and under what circumstances an individual might need additional support in order to succeed."

## Constructing Longitudinal Questionnaires

Technical issues in constructing longitudinal questionnaires are described by Endo (1992), including scales that allow room to measure improvement or decline. This discussion is beyond the purpose of this monograph, but is essential to success in longitudinal studies. "The common dilemma is whether to 'buy, build, or borrow'," explains Terenzini (2010, p. 39). Nationally recognized instruments have been developed by experts, have been field tested, have known psychometric properties, and have national norms and scores for comparison. However, "commercial measures are necessarily general and lack the specificity needed to focus in any detail on local conditions" (p. 40). Locally developed measures are, on the other hand, untested and "of unknown reliability and validity" and require faculty time and competence to develop.

Both approaches suffer from "reactivity" – "Respondents to tests and surveys know they are being studied and that knowledge may influence their responses in varying and unknown ways. Such intrusive methods influence and shape, as well as measure. Unobtrusive measures – ones that do not require a conscious response from the subject – can be highly useful as well as efficient" (Terenzini, 2010, p. 41).

Padgett et al (2010) write that "the design of the study and its constructs trump any analytical technique. Methodologically strong longitudinal studies on college student experiences with large sample sizes and extensive pretreatment measures such as the Wabash Study, the National Survey of Student Engagement (NSSE) and its precollege companion survey, the Beginning College Survey of Student Experiences; and the Higher-Education Research Institute's (HERI) CIRP Freshman Survey and its follow-up, the College Senior Survey, permit more extensive investigation into and control for selection bias" (p. 39).

## Determining "Value Added"

"Without the pretest to establish a student's level of learning at the time of matriculation, one may erroneously attribute differences on the outcomes, net of student background and demographic variables, to differential college impact by institutional type" (Siefert et al, 2010, p. 13). Including a pretest results in "statistically significant differences between institutional types on some of the posttest measures and elimination of statistically significant differences on other measures but also in substantial changes in the magnitude of the statistically adjusted differences" (p. 13). Without it, "the effects of an independent variable can be speciously over- or underestimated" (p. 5).

Terenzini explains that "change is often construed as 'value-added,' a frequently heard phrase that can be highly misleading and damaging if not understood" (2010, p. 41). "Value-added is both a metaphor and a research design. As a metaphor, it is a vivid and useful term focusing our attention on institutional effects rather than resources. Unfortunately, it can sometimes be too vivid, leading people inside and outside the academy to expect more of our assessment programs than can possibly be delivered. The reason for this lies not only in the metaphor's implication that 'change' occurs, but also that it is positive change or growth. Can 'value' be 'added' without positive change? Legislators and others are likely to say no. And therein lies the perniciousness of the metaphor, for it is important to distinguish 'change' from collegiate 'impact'"(pp. 41-42).

"Simple difference scores are highly unreliable, and they can be shown to be negatively correlated with pretest scores," explains Terenzini (2010, p. 42). "The higher the correlation between pre- and posttest measures, the lower the reliability of the gain score." Also, "simple difference scores are also subject to ceiling effects: students with high pretest scores have little room for improvement and thus are likely to show smaller gains than students with lower initial scores. Similarly, gain scores are subject to regression effects, the tendency – due strictly to measurement error – for initially high (or low) scores to move ('regress') toward the group mean upon subsequent retesting" (p. 43).

It is recommended the reader review Pascarella and Terenzini's (2005) vision and summary of *How College Affects Students* for its extensive treatment of "value-added" methodology.

## Program Evaluation Issues

While there is much interest in knowing whether "participation in a curricular or co-curricular program or activity" improves success, this is difficult to know because "students are rarely randomly assigned to programs and instead select whether or not to participate" (Padgett et al, 2010, p. 30). This means that "The researcher cannot unequivocally claim that the program effect is attributable to the program experience or to an amalgam of factors that may have influenced the student's decision to participate." This can be taken into account to some degree with the use of "propensity score methods as a means to adjust for factors that influence selection into programs" (p. 30). "Propensity score methods force researchers to clarify and examine the overlap between treated and untreated groups to ensure similarity in the likelihood of selection into the treatment" (pp. 39 - 40).

Propensity score methods use pretreatment variables that are believed to affect the program selection and that are related to the dependent variable. The treatment effect is estimated "by matching students who participated in the treatment program with observationally equivalent students who did not participate" (Padgett et al, 2010, p. 31). Padgett et al describe this methodology and use longitudinal, pre-/post-test data from the Wabash Study to examine the effect of volunteer and community service during the first year of college.

Self-selection into programs such as learning communities is examined by Pike et al (2011). Longitudinal data might suggest that participation in learning communities is associated with higher GPAs in the first semester after controlling for pre-college variables and student characteristics. "However, when instrumental variables were introduced to account for self-selection, the effects of themed learning communities on grades were not statistically significant" (p. 194).

Recognizing the design limitations placed on understanding program effectiveness, Lesik (2007) notes the impracticality of assigning students to a program such as developmental education with random assignment. A method which "embeds a regression-discontinuity design within the framework provided by discrete-time survival analysis" is presented. This provides "an unbiased estimate of the causal impact of participating in a developmental program in mathematics" (p. 583). The author was able to confirm that the risk of attrition for students participating in developmental math was "significantly lower than for equivalent students who did not participate in such programs" (p. 583).

Pascarella (2006) suggests a two-step process in analyzing the "net total effect of the intervention on the outcome," with the need to understand "the underlying processes or mechanisms" (p. 515). In a second step, "measures of the underlying processes or mechanisms hypothesized to account for the effect are added to the step one regression model. If these… actually explain or account for the impact of the intervention, then two things will occur in the step two regression model." The measures "will be significantly linked to the outcome" and "the variable representing simple exposure to the intervention, which was statistically significant on step one, will be reduced to nonsignificance on step two" (p. 515).

## Sampling

If sampling is done, it is necessary to determine whether the cohort at the outset and over time is representative of the population in terms of input, environmental, and output variables (Endo, 1992). Larger samples are required for longitudinal studies "because of the unavoidable, subject mortality problem," resulting in increased direct and indirect costs (Terenzini, 2010, p. 39).

If only a sample of the population is to be used in a cohort, for example because of the cost of an assessment instrument, then the number of respondents needed "depends upon the homogeneity of your cohort in terms of background characteristics, the number of analyses you plan to do on specific subgroups, and the number of years you intend to study these students" (Endo, 1992, pp. 28-29). Citing the work of Dillman, Bauer (2004) explains the importance of intensive efforts to communicate with study participants to get adequate response rates. She states that while "response rates for the first contact are a challenge… Longitudinal researchers have the added burden of retaining students over two or more waves of surveys." Researchers "need to disclose fully and clearly why students' responses are valued, how the data will be used, and how the respondents will receive follow-up findings" (p. 85). High response rates "can be achieved with continued contact (personal face to face is best) and reiteration of the purpose and intent of the data collection. Especially within the longitudinal design, incentives can contribute to increased continuation" (p. 85).

Thomas et al (2005) discuss problems with large-scale sample data, including the lack of a simple sampling frame from which to choose subjects and ensuring sufficient respondents with the characteristics of interest. Stratified, multi-stage cluster sampling strategies are recommended such as oversampling. The process of weighting of oversampling is described. These weights "need to be applied to the data to deemphasize the disproportionate contribution of those elements that were oversampled" (p. 57). "Adjustments for sampling error in these designs are essential for generation of accurate population estimates" (p. 71).

## Multiple Methods

Terenzini (2010) explains that a "way to ease concerns about the measurability of student progress is to ensure that multiple measures are incorporated into one's assessment program" (p. 35). "Multitrait, multimethod matrices can be highly useful" and incorporate triangulation. "Adoption of multiple measures is likely to have more face validity that will appeal to faculty members as well as increase the confidence that can be placed in interpretations of the data" (p. 35).

As mentioned, Reynolds et al (2010) use a mixed methods approach with qualitative interviews to enhance the longitudinal depiction of persistence and success. A qualitative approach can provide a rich dimension to the results, with narrative that helps bring the data to life. The reader is referred to Howard's (2007) edited AIR monograph *Mixed Methods in Institutional Research*. Qualitative methods are useful in identifying themes and possible patterns in the data, for providing a story-telling type narrative. Yet the user can't make any type of quantitative conclusions and must be extremely cautious in attributing any type of meaning to even simple descriptive statistics such as frequencies and distributions.

## Section 8. The Technology of Display

Achieving the Dream (ATD) describes the process of "Building Institutional Capacity for Data-Informed Decision-Making" and the barriers that may exist. One common factor seen in the analysis of LDS efforts is that longitudinal results are not "tailored for specific thus, finding information even in a centralized system can require technical skills possessed by only a few individuals" (ATD, 201?, p. 6). "High-level and grassroots demand for the system is critical" (NFES, 2010a, p. 10) and this requires "timely, useful data for a variety of stakeholders" to "access, analyze, and interpret the results" (p. 11).

Technology must be leveraged to display results so that customized information is pushed to the desktop, iPad, tablet, and smartphone; however, the user wants to use them. Currently, users have a high level of expectation that software applications will give them personalization and drill-down capability. Online tools to do this are presented as dashboards, attention management systems, and scorecards. How will these be leveraged to display the results of longitudinal studies? This section provides a brief description of the expectations that exist now for dynamic presentation of data in higher education. The types of new business intelligence tools that are possible are reviewed, including how they might be used for longitudinal purposes. Some of these tools come in an environment that is quite costly. Therefore, a brief focus is given to open source or low cost tools that are available or are within the existing suite of software suites such as Microsoft SQL Server but perhaps less known. Finally, several pragmatic observations about the IT change process are put forward, such as when and how to compromise on the availability of tools and selection of a database.

### Dynamic Presentation Expectations

The issues in presenting tables and graphs of longitudinal data are essentially the same as in presenting other types of IR studies. Effective presentation requires effort and a nuanced understanding of data-infused decision-making and the "politics of data." This is the topic of another AIR *Handbook of Institutional Research* chapter, other AIR publications, and many other resources.

The advent of the web and expectations for the use of online tools draw upon the On-line Analytical Processing (OLAP) model with customization of reports based on the needs and interests of the user and the ability to drill down or up on various levels of aggregation. These features provide the ability to move seamlessly through customizable tables and graphs of data, clicking and selecting points of interest to focus on what's important. Anticipating all of these custom levels of reporting and aggregation is the work of good software designers and good report writers and the need for these skills is just as great for presenting the results of longitudinal studies as for other types of IR data.

The nature of longitudinal displays is quite complex, requiring extensive preparation time to update the data with the latest Clearinghouse and employment records to accommodate all types of enrollment, transfer, completion, industry licensure/certification, and employment outcomes. These are the same questions being raised across the country with SLDS grants to states. What is the most interesting and meaningful way to present accurate, timely, and relevant longitudinal data that will attract the interest of and engage different types of stakeholders to use these data for informed decision-making?

Longitudinal reports would run hundreds of thousands of pages if printed, fill millions of spreadsheet cells with data on persistence and success rates by program for different subpopulations of a cohort over time. As valuable as these static reports of data might be, they are worthless to users unless they can customize it to the specific cohort, time period, and variables of interest; unless the data can be stored in a format which will allow for further analysis, such as SAS or SPSS; and unless the results can be viewed at different summary levels of aggregation – at the macro level for the institution as a whole and at the micro-level (i.e. a specific full-time, first-time cohort of transfer majors in a discipline tracked over ten years into the workforce).

## New Business Intelligence Tools

New business intelligence and analytics tools for higher education are being leveraged extensively, utilizing student, course, award, finance, personnel, and workforce data in data marts and data warehouses. The reader is referred to the recent *Handbook of Institutional Research* for a chapter about this topic (Milam et al, 2012). It is important to gain perspective on what is possible and where the future appears to be heading in the presentation of longitudinal data.

Longitudinal data may be part of a data warehouse, data mart, or online data application. All of these applications store tables of data in a database software package such as Microsoft SQL Server, Oracle, and/or MySQL. This database structure may be standalone, not part of a larger online solution, but simply residing on the user's desktop. Where there may be an infrastructure of reports available in the data warehouse or data mart, the users of standalone databases are typically left to their own devices in writing SQL and using WYSIWYG[12] tools to create reports and visual displays of results.

Online vendor reviews and products such as BI Scorecard document the features of software regarding the levels of customization, display, and aggregation provided (Howson, 2012). It is the underlying database technology of cubes and dimensions that makes this possible. It is recognized that the IT support at the institution will for the most part drive the choice of database, operating system platform, and BI visualization tools. Some of these choices are quite costly and others are open source. The user needs to ask whether the results of a longitudinal study are expected to be presented on a mobile device such as an iPad? If yes, then the institution should already be experimenting with mobile solutions and be BI-savvy; or the IR staff need to be

[12] What you see is what you get.

personally immersed in leveraging this technology. The solution of how best to use business intelligence and analytics will, very often, not be handed to the researcher if it is not already in place.

The Gartner Group provides an analysis of BI tools in its "Magic Quadrant for Business Intelligence Platforms" (Hagerty et al, 2012). Developers of longitudinal studies need to address the BI infrastructure, how they will manage metadata using the tools provided with the CEDS and other resources, decide upon what development tools they will use, and collaborate with other BI users on content, data structures, metrics, and other issues of what is now being called master data management (MDM). Longitudinal data must be seen as one more wave of the "exaflood" of big data that have come to be in the past decade (Fishman, 2010). "Big data" are so complex, unwieldy, and different in structure and content that MDM strategies are necessary if data professionals in the 21st century are to have any hope of keeping up with demand and expectations for display and use. Longitudinal data structures and techniques such as data mining and exploration are critical for success in the complex expectations of the higher education community for addressing big data, policy issues.

In terms of reporting, users expect to be able to "create formatted and interactive reports, with or without parameters, with highly scalable distribution and scheduling capabilities" so that people can "access and fully interact with BI content delivered consistently across delivery platforms including the web, mobile devices and common portal environments" (Hagerty et al, 2012, p. 1). Often visualized as dashboards, BI tools must have "intuitive interactive displays of information, including dials, gauges, sliders, check boxes and traffic lights. These displays indicate the state of the performance metric compared with a goal or target value. Increasingly, dashboards are used to disseminate real-time data from operational applications or in conjunction with a complex event processing engine" (p. 1).

At times, users need to be able to "ask their own questions of the data," with a "robust semantic layer" that allows ad hoc queries (Hagerty et al, 2012, p. 2). BI products should be integrated with the office tools that users rely on, primarily Microsoft Office, with "support for document and presentation formats, formulas, data 'refreshes' and pivot tables." Structured and unstructured data sources should be available, with search features that map them into "a classification structure of dimensions and measures." Mobile BI tools for tablets and smartphones should take "advantage of the interaction mode of the device (tapping, swiping and so on) and other capabilities not commonly available on desktops and laptops, such as location awareness" (p. 2). OLAP allows "a style of analysis known as 'slicing and dicing.' Users are (often) able to easily navigate multidimensional drill paths" (p. 2).

With interactive visualization, BI tools bring the "ability to display numerous aspects of the data more efficiently by using interactive pictures and charts, instead of rows and columns. Over time, advanced visualization goes beyond just slicing and dicing data to include more process-driven BI projects, allowing all stakeholders to better understand the workflow through a visual representation" (Hagerty et al, 2012, p. 2). Predictive modeling and data mining are possible and models are incorporated into BI reports and dashboards. BI takes the "metrics displayed in a dashboard a step further by applying them to a strategy map that aligns key performance indicators (KPIs) with a strategic objective" (p. 2).

With tools such as Visual Tableau, MicroStrategy, JasperForge, OmniGraffle for the Mac, and various Adobe products, the user does not need to know ahead of time how she/her wants to display and drill down on longitudinal data. These tools recognize two types of variables when data are imported into the system – categorical variables or dimensions and continuous variables or measures; and they allow measures to be quickly converted into dimensions and vice versa. Special toolsets are provided for working with geographic data, with mapping overlaps based on existing Geographic Information System (GIS) data and products. Once data are brought into these tools, the software automatically shows the different types of visualizations that are possible.

The possibilities for visualization are endless and the software is capable of doing many things – albeit some of them not particularly meaningful or interesting for understanding SLDS. BI tools may be focused on dashboards and their creation and dissemination (such as iDashboard) or on a broader approach to data visualization that integrates dashboards as one mode (such as Visual Tableau and MicroStrategy). For these tools, data may be imported into proprietary storage formats or maintained in their native SAS, SQL Server, Oracle, or other database package. The key lies in Open Database Connectivity (ODBC) connections between BI tools and different proprietary databases, passing user credentials in order to obtain different types and levels of access and query capability. This same approach is possible with spreadsheets set up to access Student Information Systems extracts, as long as permissions and connections are in place. Login userid and password credentials are typically maintained with LDAP[13] and other directory permissions.

### Open Source Alternatives

Open source tools, where the underlying coding or programming is made available to the world and contributors may improve and extend the functionality, are an attractive alternative. Open source does not necessarily mean free, however. Still, "nearly all free software is open source, and nearly all open source software is free" (Free Software Foundation, 2012). Some of the free, open source BI tools currently available are Eclipse BIRT, RapidMiner, SpagoBI, R, KNIME, Pentaho, Jedox, and JasperForge. SourceForge is the standard online source for open source software; see http://www.sourceforge.net. A list of these is provided on Wikipedia. Other free but proprietary tools are available, though these are sometimes developed as a way for vendors to get their "foot in the door,"

[13] The Lightweight Directory Access Protocol (**LDAP**) is an Internet protocol for accessing distributed directory services.

with free products that unexpectedly morph into expensive licensing solutions. Two, currently free products are InetSoft and QlikTech (Wikipedia, 2012). The open source product programing language R is recommended by King et al (2012) in their analysis of complex data for IR as being "extraordinarily flexible" and supporting "a vast array of statistical models."

In terms of database software, popular alternatives include PostgreSQL, MySQL, Apache Derby, MSQLDB, and H2. While spreadsheets are discussed for their use in BI, the free tool PowerPivot is available from Microsoft. This is bundled with SQL Server and for working with Microsoft's SharePoint environment. PowerPivot for Excel and for SQL Server provide a WYSIWYG environment that takes spreadsheet Pivot Table functions to a much higher level. Where it is fairly difficult to join data across multiple tables in Excel, PowerPivot allows much more database-type functionality. Visual diagrams are used to link and manage data. The BI tools allows the user to "bring data from virtually anywhere; Easily organize, connect, and manipulate tables of large data sets; Perform in-depth analysis of your data, any way you want to slice it; and Use PowerPivot for SharePoint to share your workbooks across your team or publish them to the Web" (Microsoft, 2012a, p. 1). The SQL Server 2012 release of PowerPivot provides key performance indicators and an Indicator Dialog Box. Both versions have series of special data functions (called DAX) that provide a data manipulation language with statistical and information functions. A Wiki is provided on the Microsoft PowerPivot website with information, downloads, and links.

If Microsoft SQL Server is owned, there are extensive tools for BI within this software that may be used, including Report Builder and Power View. With Report Builder, features such as row and column groups, sparklines, indicators, graphics, and maps may be displayed in a report. The user specifies "where to get the data, which data to get, and how to display the data," then publish the report to a report server or SharePoint node (Microsoft, 2012b, p. 1). The Power View Add-in for SharePoint provides "an interactive data exploration, visualization, and presentation experience," with "intuitive ad-hoc reporting" and the ability to "easily create and interact with views of data from models based on PowerPivot workbooks." The environment is "similar to Microsoft Office" with the menu ribbon and the ability to move data between different visualizations (Microsoft, 2012c, p. 2). Oracle, SAS, and other vendors all have their own equivalent BI toolsets. These are detailed in the Gartner Magic Quadrant report with the analysis of strengths and cautions (Hagerty et al, 2012).

The new version of Excel 2013 includes a number of BI tools and incorporate the database features and DAX programming language in PowerPivot. These BI tools are all being moved to the Cloud, so there is less need to install software on the user's local machine. With a single point of login, BI applications and data are available through the Internet.

### OBSERVATIONS FOR THE IT CHANGE PROCESS

In promoting any kind of data-informed approach, it is unfortunate that the availability of software sometimes drives the analysis, not the other way around. It should be recognized that longitudinal studies are not an IT solution, but one tied to emerging interest in BI and analytics (much of what is still called IR).

If a dashboard tool is already being used, then it can be leveraged for longitudinal data if the data can be stored in the needed dimensions and table structures and updated regularly. If no such tool is available, the presentation of longitudinal data may be the place to start, but there is already a steep learning curve involved in the ETL and development of the unique, institutional data structure. Learning another set of tools may be too much effort and time. These tools are moving away from the need for syntax-level programming, replacing it with visual, graphic tools and embedding modules and functions that can be called up at any time. However, there is a quickly expanding knowledge base and set of experienced state-level staff looking at longitudinal data, funded by SLDS and mandated by states' accepting ARRA money. Decisions about IT and staffing and project priorities are being made now and very quickly before time and money run out. Therefore, several pragmatic observations are put forward for consideration by the reader.

First, the choice of a database doesn't really matter – as long as it can create and store OLAP cubes with standard dimensions at different levels of granularity that can then be used with data visualization tools such as dashboards on desktops, tablets, iPads, smartphones, and in the Cloud.

Second, the software package for the data warehouse or data mart doesn't matter – as long as it allows for the extraction and transformation of clean, reliable data. Some form of SQL (syntax or visually) is used to generate customized reports, merging tables and creating derived variables and using taxonomies with different levels of aggregation. Ideally, this process will be Cloud-based because it is ubiquitous, equipment costs are lower, software may then be priced based on usage and demand, and it is tied to existing security protocols. The data structure of the applications needs to allow for SQL that dynamically passes parameters to a database and returns results so that users can drill up and down on cubes of data documented as measures and dimensions. In buying a vendor product in this area, you are really purchasing the vendor's proprietary knowledge of your data structure in student, finance, financial aid, and human resource information systems.

Third, the BI visualization software doesn't matter – as long as it allows teams to create custom displays, is not wedded to a specific data structure, and helps users discover new patterns visually in complex data. You may want to use the evolving BI capabilities of Excel as a start and document your assumptions about what users want and need.

Fourth - what matters most is that a solution be put in place and used. The solution can't be so unwieldy in politics, staffing, IT infrastructure, or cost that it cannot be learned and managed. All of these tools are evolving and changing to meet the growing expectations and emerging requirements for dynamic data presentation.

The technology for linking different sources of data is explored by Diaz (2010) for a SHEEO presentation. These dimensions of technology and data usage are arrayed to describe four types of settings for linked longitudinal data – high and low technology versus high and low usage of multi-source data.

- High technology - "Regularly scheduled, comprehensive linkages within and between agency data systems. Easy access to timely data. Tiered access for stakeholders."
- Low technology - "Data in unlinked silos. Individual, manual file transfers on request or with approval. Data access is limited to those with significant expertise or permissions. Data sharing is ad hoc and not automated."
- High usage – "Usage of data is integral to agency mission and employee work plans. Usage is frequent and regular. Data are easily requested, while still protecting privacy. Users are taught how to interpret data. Users' feedback is sought to improve data products. Guidance is provided on what interventions might follow from the results of an analysis."
- Low usage – "Usage of linked data is occasional and irregular. It is complicated or time-consuming to request data. Data are reported, but no one is sure how they are being used – there is no feedback look between providers and users" (Diaz, 2010, p. 6).

If the analysts are doing their job, the heart of the issue is the extraction, transformation, and loading (ETL) of data. Next come the data administration tasks necessary to ensure clean, reliable data and naming conventions. These data management processes will ensure the ability to merge, recode, and handle complex data through cubes, dimensions, measures, and taxonomies. If these things are all in place, then whatever tools are available can be leveraged successfully to analyze longitudinal data.

## Section 9. Sources of Data

Much of the focus of this monograph has been on enrollment, course, completions, and employment data. To understand the range of topics that may be addressed with longitudinal studies across the continuum, other types of data need to be incorporated.

### P-12 Schooling

The use of early learning data for longitudinal studies is discussed by Jenner (2009) of Washington's Education Research and Data Center (ERDC). These include Washington State-supported early learning programs, Head Start, and state-funded child care. The use of these data was spearheaded by Washington Learns, a review of the state's education system which resulted in a "teacher-course-student collection" and the creation of the ERDC. Some of the PK-12 data elements brought into longitudinal studies include: birthdate, gender, race/ethnicity, grade, school, expected graduation date, enrollment and exit dates, free- and reduced-price lunch eligibility, participation in programs such as Title I, migrant status, homeless status, language spoken at home, test scores, course-taking behavior, credits, and grades.

Prescott and Ewell (2009) add several more interesting K-12 elements, including upper-level math course, upper-level science course, AP course, and state exam score. Numerous K-12 examples of longitudinal studies are presented in the National Forum on Education Statistics Guide series. One topic, the effect of high school requirements on persistence, is explored by NCES and others. Four high school diploma options in Indiana are used to understand student performance. In addition to student-level data from the Indiana Commission for Higher Education and the Indiana financial aid agency, data were brought in from the LearnMore Resource Center survey of ninth graders and the SAT Student Descriptive Questionnaire (Musoba et al, 2008).

With longitudinal, SUR data, researchers have been able to "control for student background characteristics, educational aspirations, institutional characteristics, and socioeconomic status in assessing the effects of increased high school graduation requirements on student academic success (Musoba et al, 2008, p. 99). The results help administrators understand the impact of raising graduation requirements on high school graduation rates, enrollment, and early persistence. College choice models may then be explored in the context of diploma type. Furthermore, these data allow users to "model the complexity of the interactions of race/ethnicity, income, academic preparation, and financial aid to address important questions about educational access and equity" (p. 99).

A Teacher Certification Data Base in Missouri is discussed by Wittstruck et al (2002) that contains information on all teacher certificates by type issued since 1970. These data may be used to help document teacher quality and may be linked to other sources to track teacher earnings. Both types of data may be tied to the ACT score ranges of students and to the percent of students eligible for free and reduced lunch programs to evaluate performance.

Linking inter-agency K-12, teacher certification, postsecondary enrollment, and credential data, a number of key questions may be addressed. For example, "What programs are producing college ready graduates?" "Which teachers' students have unusually high college and workforce success rates?" "What high school performance indicators predict success later?" and "How does dual enrollment influence student success?" (Diaz, 2010, p. 4).

A longitudinal, teacher pipeline application is described by McLaughlin et al (2008) that contains data on K-12 teacher candidate enrollment, program descriptions, licensure and school employment, follow-up surveys, employment, school working conditions, and learner performance. The Virginia Improves Teaching and Learning (VITAL) data system was built for the State Council of Higher Education for Virginia. VITAL was designed to collect and store individual SUR data in a Teacher Education and Licensure (TEAL) data warehouse. The system provided multiple ways to track students during and after teacher preparation programs. Users create custom extracts with syntax programs that output the data in multiple formats, such as SAS and SPSS. Unfortunately, state funding priorities did not allow Virginia to maintain this system for more than a few years after launch.

In tying P-12 data to workforce, "Perhaps the biggest challenge is matching K-12 students who are not found in one of the states' postsecondary databases with available employment records if the K-12 system does not collect student Social Security numbers or is prohibited from doing so" (Prescott and Ewell, 2009, p. 13). Alternative approaches to identity matching are then required, using combinations of fields, unless a unique, statewide student identifier is in place.

Expanded LDS features are documented by the National Forum for Education Statistics' LDS Guides. For K-12 LDS, these include data on interim and formative assessment, finance, facilities, and geographic information (NFES, 2010a). The types of P-12 questions raised by the LDS Guides are ones that postsecondary SLDS and institutional models might also address, depending upon the context. These include questions about groups of students; teachers; policies, programs, and strategies; indicators; schools; districts; and states.

## Social Services

Jenner (2009) describes the use of TANF, Basic Food (food stamp), Foster Care, and Medicaid data from social services as part of a larger continuum of movement through the social welfare system and understanding success. Data sharing agreements were put in place through the Washington ERDC, along with "data-linking strategies anonymization protocols." In support of TANF administration, agency data were linked to confirm low-income status and citizenship and the data were used for "WorkFirst program administration."

Colleges do not usually collect information about welfare recipients, relying instead on Pell eligibility from data on the FASFA and residence in economically disadvantaged areas as proxy measures for low-income status. FASFA data are not available for many students, so this data point can be incomplete. Some states such as California and Washington are able to match SUR data with TANF records (Jenkins, 2008). Data may also be linked to a state Vocational Rehabilitation Data Base and Vocational Education Database, such as used in Missouri (Wittstruck et al, 2002).

One example of the longitudinal use of social service data is the Great Expectations program of the Virginia Community College System. Fueled by the recognition that foster care youth have postsecondary enrollment at a far lower rate than those outside of the foster care system, this federally and state-funded program provides intensive supports for these students to help ensure success. Using longitudinal data, it is possible to see whether the program has an impact; but only if social services data to identify students previously in foster care are leveraged. Foster care is not documented in admissions applications or the FASFA and only limited numbers are served by the special scholarship programs provided. Self-identification can have its own therapeutic downside. Similar efforts are being made, with state variation in their interpretation, as part of adoption family preservation grants, partly through scholarships for students who were special needs adoptions or in foster care. FERPA protections and program-specific restrictions on the release and use of these data need to be carefully maintained. Yet there is a middle ground for longitudinal research to serve the greater good.

Programs such as Pathways out of Poverty are designed to meet the needs of low-income students and give them the skills and income so that they no longer need food stamps, Medicaid, WIC, and other services and can become employed and self-sustaining. These programs are of great political and social interest, especially with federal and state budget constraints. If higher education is to be seen as a springboard for human capital and improving the economy, then social services data need to be leveraged in longitudinal studies. With the increased visibility given to states' focus on increasing enrollment and credentials, LDS data are becoming more available. Examples of longitudinal studies promulgated by Jobs for the Future and the National Governor's Association need to be disseminated as best practices for using these data.

## Financial Aid

Financial aid data include types and sources of aid, income, amount received, need status, and debt (Pai et al, 2008). "State and institutional financial aid efficacy is best determined while controlling for academic preparation because academic preparation is often offered as an explanation for the low-persistence rates of low-income students," explain Musoba et al (2008, p. 101). The authors find that aid is positively associated with persistence in the first year, but negatively from the first to the second year. First-generation students' grant aid is positively associated with persistence through the second year. Aid is associated with Hispanic and African-American student persistence, but not for white students (Musoba et al, 2008).

Research on student loan defaults is described by Wittstruck et al (2002), focused on "determining characteristics of students who are more likely to default on their student loans" (p. 10). The results suggest that "Students who remain continuously enrolled in college, eventually complete degrees, and have a lower debt burden upon exiting college are less likely to default on their student loans" (p. 10). Default rates are explored by degree level, continuous enrollment periods, annual earnings, and debt load in analyses. In Indiana, the impact of financial aid packages on persistence for different student populations was tracked, including issues related to first-generation students (Musoba et al, 2008).

The CMSS notes that HEOA requires two employment measures tied to Social Security earnings and student loan debt: debt-to-earnings ratio and loan repayment rate. Under the auspices of GE reporting, there should be "useful insights into the employment outcomes of program completers at two-year institutions" (CMSS, 2012, p. 25).

There is a significant amount of scholarly research using longitudinal data on financial aid. An example is a study by

Fenske et al (2000) of four cohorts with breakouts of women, minority and low-income students by major (STEM versus non-STEM) and the impact of factors such as financial aid on persistence. STEM majors took longer to complete, but did so at a higher rate. "Women, underrepresented minorities, and needy students received more gift aid than other student populations" (p. 67). Gift aid for STEM majors was "more likely to be based on merit rather than need." Loan debt increased rapidly for both STEM and non-STEM.

Three institutions in one of the IPAS projects were concerned about student employment and its consequences on academic success. One thousand students at each campus were surveyed and the data merged with student demographic, financial aid, and persistence data to "develop a descriptive picture of the relationships between employment and persistence… as well as model these relationships while controlling for other factors known to affect retention" (Musoba et al, 2008, p. 104).

### EMPLOYMENT

The Voluntary Framework of Accountability (VFA) measures include the percent of students completing a program or 90 contact hours that are employed with a livable wage and median wage growth. A living wage is set at 200% of the poverty level for a family of four. The VFA documentation states that Unemployment Insurance (UI) wage data should be examined five quarters after completion. The calculation of wage growth excludes those making below minimum wage for a full-time quarter of work and students who are still enrolled in higher education within five years (AACC, 2012).

Employment records may be used to "construct measures of job placement and earnings by field of study and/or field of employment" (Ewell and L'Orange, 2009, p. 7). These data help "determine how long it takes graduates or former students to gain employment and the individual return on investment associated with further education. Finally, it can help students determine their odds of gaining successful employment in the field(s) of their choice" (p. 7).

DQC (2010) documented the status of states' ability to make "Education and Workforce Data Connections." Once in place, these become the heart of an "Integrated State Workforce-Postsecondary Data System" (Phillips, 2009). Workforce Investment Act (WIA) data are included in this system. "Wage outcomes are a compelling accountability measure," explains Schenk (2011). They "indicate whether students receive financial gain and are also an indication of performance in the workforce" (p. 2).

In calculating these indicators, it is useful to differentiate them by "different departure conditions (e.g. graduated/not graduated/number of postsecondary credits earned by time of departure" (Prescott and Ewell, 2009, p. 9). This is particularly helpful in documenting the success of community college programs where the vast majority of students in a major/plan do not finish the associate's degree, but do go on to employment in a related field. Employment salaries should be calculated for those that do and do not complete, looking at wages in the calendar year before and after year-of-enrollment and before and after year-of-graduation.

Job placement rates are a longitudinal measure promoted in the Completion Arch, but "relatively few states report employment statistics for community college students" (2012, p. 55). "Students with earnings in a given time period are counted as employed, and students with no earnings are counted as not employed." The UI wage data do not include "individuals who are self-employed, employed by the federal government (including the military) or employed out of state" since they do not "pay into a state's UI system" (p. 55). This underestimates the employment measure. The data also do not exclude people who are "not looking for work" because of education and family reasons. Virginia is touted as a state with "comprehensive indicators of job placement." Some of these state reports do not, however, distinguish students by whether they earned a certificate, degree, or no credential. Wyoming and Wisconsin "use surveys of former students to measure their employment outcomes," but it is noted that these may not be representative of the student population.

National BPS longitudinal data are used by NCES to calculate employment rates for the 2003-04 student sample as outcomes six years after first enrollment. These may be broken out by award level, field of study, and demographics. The BPS rates are higher than those estimated by states such as Virginia which are based on UI wage data, in part because the BPS uses student interviews. The data suggest that students "who completed a certificate or associate degree were more likely to be employed than students who did not" (Completion Arch, 2012, p. 56). Students in STEM fields earned the highest salaries. Where nontraditional students often have "less favorable outcomes," the BPS data suggest that "students who started community college at age 24 or older earned appreciably more on average than younger students" (p. 58).

Mean and median wages after employment may be studied with these data. States like Illinois "report students' wages before and after attending a community college" and "the difference can be interpreted as a rough estimate of the effect of attending a community college on students' earnings" (Completion Arch, 2012, p. 57). Unfortunately, "there is no common standard" for reporting estimates of wages. "Wages may be reported as hourly, quarterly or annual earnings" requiring conversion to a standard. While "a few states report wage data longitudinally," most "limit estimates to one point in time" (p. 57). Oklahoma calculates averages salaries at one year and five years after graduation from a four-year, public institution. Wage records in Oklahoma are supplemented with those from Tax Commission data, which "cover many federal employees and self-employed individuals who are not included in UI data" (p. 57).

The State Council of Higher Education for Virginia (SCHEV) recently released the workforce component of the state's forthcoming SLDS (called the Virginia Longitudinal Data System) and it includes wage data by program. A number of caveats are put forward and the user of these data is cautioned about their interpretation; yet they bring to life the essence of what a SLDS can do. Data are presented by CIP code, award level, and institution for those programs where there are sufficient numbers of graduates and UI wage record data. No one year of data is presented; rather, rolling, five-year averages are shown. There must be 10 or more full-time equivalent graduates to be reported and at least 30% of a program's graduates must have wage data, 20% if there are 200 or more graduates. Not all programs are included, since they must have produced at least three graduates in each of the five, successive years. Wages are not adjusted for inflation. Wages are reported as of 18 months after graduation. SCHEV's web-based portal provides access to the data with a wide range of reports. In addition to mean and median wages, the percent of graduates that are in the workforce and of those still enrolled in postsecondary education is reported.

In working with wage data, it should be recognized that employment rates are "strongly influenced by local and regional economic conditions that are largely beyond the control of community colleges" (Completion Arch, 2012, p. 55). In providing "A Short Guide to 'Tipping Point' Analyses of Community College Student Labor Market Outcomes," Jenkins (2008) describes the use of UI wage records for eight quarters before the first enrollment to the latest available in order to "examine the effects of milestone attainment on changes in earnings independent of other factors that have been shown to be correlated with student success" (p. 1).

Student engagement is tied to early career earnings after college by Hu and Wolniak (2010). Analyzing data on freshmen collected with three waves of longitudinal surveys from the Gates Millennium Scholars program, the authors look at student engagement that occurs in college activities and link it to employment data and major (STEM versus non-STEM). The results suggest that academic engagement may be "positively related to early career earnings for non-STEM students" but negatively related for STEM majors. Social engagement was positively associated with career earnings for STEM graduates but not non-STEM. There is a "complex relationship between student engagement and early career earnings" (p. 750).

In "Measuring Transitions into the Workforce as a Form of Accountability," Schenk (2011) describes using SUR data to "follow students into the workforce." Data on completions and employment are tracked for a cohort of community college completers and "leavers" in the 2005-06 academic year in Iowa.

Data on majors and industry of employment are analyzed, with a "unique graphing technique to display the relationship between majors and industry." These graphs are "an effective tool to communicate results to various stakeholders" (Schenk, 2011, p. 2). The typical display of tabular data does not allow the user to "identify significant points of transition. Readers are essentially left to find the biggest number in the table in order to see the most significant results" (p. 4). It is not very easy to convey a sense of proportion and it is difficult to "understand the origin major from the industry perspective." With this analysis and type of graphical display, the communication of results is improved with a wheel of colors and threads that visualizes the flow between clusters of majors and industries. Note that questions such as "are there enough majors going into an industry?" are not answered (p. 9). This is not the purpose of the project and while interesting, it requires further analysis. The lack of occupational codes in most UI wage data is a problem in pursuing this topic.

The analysis of CIP to North American Industry Classification System (NAICS) requires a crosswalk from CIP to SOC and from SOC to NAICS. With each step of this conversion, the level of granularity and the accuracy diminish. However, since UI wage data do not routinely include SOC data, only NAICS codes, multiple crosswalk translations are needed – from CIP to SOC and from SOC to NAICS. The reader is warned to expect vagaries in how a CIP to NAICS crosswalk may be applied and used. The one-to-many and many-to-one relationships are problematic and must be addressed in a way that is sensitive to the issues of interest.

Acknowledging that many schools may not have access to employment data, the CMSS encourages ED to give incentives for "more robust data systems" and "increased guidance on data matching." Schools need assistance in the interpretation and disclosure of employment rates. This was done to some extent with GE reporting requirements. Schools were required to document the SOC codes for which its GE CIP programs provided preparation. Non-GE programs do not fall under this scrutiny. It is hoped that SLDS grants will result in "systems that make employment data easier to access and use that increase the capacity for interstate sharing" (CMSS, 2012, p. 26).

### Industry Licensure/Certification

Adelman helped the higher education community understand that there is a "parallel universe of postsecondary credentials" in industry certifications that are awarded outside of the higher education environment but as a result of its preparation (2000, p. 1). These industry licensures and certifications need to be included in the longitudinal tracking of student success if our understanding of movement across the continuum from noncredit to workforce is to be complete.

The availability of data on industry certification and licensure varies "by state and by occupation," however (Completion Arch, 2012). Texas "provides more detailed data on licensure exam pass rates" with "results for approved technical associate degree and certificate programs that lead to professions requiring licensure or certification" (p. 54). "Texas community college students taking licensure and certification exams passed at a rate of 91 percent," with variation by field. While there are extensive data on licensure rates to monitor success by industry and occupation, sometimes there are "relatively few test-takers per exam per year," so the results can't be broken out by subgroups (p. 54).

For institutions whose mission includes technical and career training, the calculation of graduation and student success rates can be more complex. As AIR (2010) explains, "The Carl D. Perkins Career and Technical Education Act of 2006 allows colleges to include students who earn industry-recognized certificates or credentials as completers when calculating the student success metrics. It can be difficult to find out which students earn this type of certification. Including these awards in your longitudinal analysis can help you to better explain student success" (p. 40). At the same time, this muddies the understanding of success in CTE programs of first-time, degree-seeking students. The same thing happens with the award of general education certificates to degree-seeking students. This increases the calculated graduation rate, but at the cost of making the rate meaningless for its original SRK purpose.

Career Readiness Certification is another type of credential offered by some community college systems. This involves an external exam, the results of which are not necessarily stored in the student information system along with other types of awards. Therefore, these data will need to be extracted and stored in special tables to supplement other types of awards. Sometimes, these certifications are done with credit courses, other times for non-credit, usually depending upon the state's approach to non-credit instruction. If the data are stored in different collection or transaction systems, the user must be prepared for problems in the use of different identifiers, calendar years, application forms, minimum datasets, recording of demographics, and missing data.

### Workforce Training and Non-credit Instruction

In a national study of noncredit instruction funded by Lumina Foundation for Education, Milam (2005) found that there is "no national statistical portrait of the impact of noncredit classes" and no uniform reporting of noncredit data across federal and state agencies. In a subsequent paper entitled *The Hidden College*, Voorhees and Milam (2005, p. 1) determined that "A very large slice of the learning marketplace operates beyond the view of public policy. Here, we refer to noncredit programs operating under the aegis of traditional higher education institutions. These programs purportedly serve millions of learners each year, but no one knows their full scope. No national data exists that traces the types of programs that attract learners nor what that volume may be."

Another, extensive review of noncredit data was conducted for an NPEC Research and Development Background Paper (Milam, 2008). The paper was completed just days before HEOA was signed into law and NCES had to address many other needs for IPEDS that were of higher priority. The conclusions of this review are that definitions of noncredit vary by state and that the flow of information for Department of Labor, WIA, and other funding streams through higher education for workforce development, noncredit, Adult Basic Education, GED classes, and English as a Second Language remediation is not tracked anywhere centrally. It is possible to collect all of the data required for these different funding streams in one place, but only at different levels of aggregation. This is usually at the institution level, but sometimes it is limited to the system, workforce investment board, Congressional district, and state.

Part of the problem is that noncredit, continuing education, and workforce data are often collected and stored using third-party (or shadow) systems that are not directly part of or even connected to student, human resource, or financial information systems at an institution. A common student identifier may be used, but it is not always collected – for example with contract classes taught at a business where the employees do not register individually but as a group. A minimum dataset is only collected for credential and contract paperwork in this case. Where data are collected, the process involves duplicate data entry in two parallel systems. This is fraught with problems, since the second data entry is driven from reports that are produced from the first, potentially proliferating errors. Another problem occurs because the calendar of federal reporting from July 1 to June 30 cannot be easily overlaid to semesters or annual, calendar year reporting. This is exacerbated with differential start and stop dates for noncredit and workforce classes that are customized to meet specific business needs and timing. Also, noncredit, workforce enrollment counts are based on completers, not those seated on the first day of class or on a census date.

When incorporating workforce and noncredit data into longitudinal studies to document the continuum of educational experience, it is important to locate data dictionaries and to experiment with the queries that are available to extract data from production and query systems. The typical approach to ETL may be inadequate, due to the duplication of students and the flow of records created at multiple points in time with different effective dates to document registration. The completion of courses and then the subsequent awarding of credentials also create separate and distinct record structures that must be extracted, collapsed, and aggregated for reporting. The documentation of credentials upon completion may be stored in another third-party system than that used for enrollment, requiring separate extracts and merges. Some of these credentials, such as licensure exams, may never become accessible to institutions as a data collection. In this case, these data must be entered manually and individually from official and unofficial paper copies retrieved from students.

Another problem occurs with the sheer complexity of workforce offerings. ABE, ESL, and GED may or may not be found under the umbrella of P-12 with its unique reporting requirements from the respective funding streams. The data might be shared as part of an inter-agency memorandum of understanding, but more likely not. Different student identifiers may be used, with the need to map these to SSNs or others in the LDS.

Community education courses may be scheduled, but without formal registration paperwork and reports generated based on counts of participants at events. The same course may be both for credit and for noncredit students, for example depending on the need for CEUs. The noncredit class data may not contain the same variables of interest collected for credit. The data may not contain basic demographics, once thought to be required to comply with federal reporting for the identification of illegal immigrants. Now that this is not always the case, depending upon the state policy environment, the use of a minimum dataset for third-party contracts in workforce can be even more of a barrier to longitudinal student tracking across the continuum.

A different productivity measure is used than the full-time equivalent (FTE) for credit. Contact hour counts are based on various criteria and time periods. Employers served is another data collection of interest to workforce managers. These data too are minimal and incomplete, focused on required reporting about the type of employer and industry and the number of individuals served. Examples include apprenticeship-related instruction, the use of career coaches, and special efforts to serve 18 to 24 year olds and help them earn a GED and move on to college and the workforce, such as the VCCS "Middle College" program. These transition services are provided to at-risk populations. Data on these educational experiences are difficult to document, extract, standardize in format, and merge with credit enrollment, completions, and employment data in order to paint a complete, longitudinal picture of student outcomes and success for at-risk youth. While complex, this is the nature of these offerings and this important work needs to be captured and shared.

### Learning Management Systems

Data on online courses captured as part of learning management systems (LMS) such as BlackBoard and Moodle provide another dimension to the continuum of educational activity and should not be overlooked. One way to view them is as another source of information about courses. In addition to instructor type, discipline, productivity, delivery mode, location, and similar data elements, there are data in LMS about the number and type of discussion threads, postings, announcements, available documents, student accesses, graded items, disk utilization, available tests, and available assignments. Analytics about these data are provided by LMS vendors in new BI tools and these help institutions track the depth and type of interaction with online courses, as well as emerging measures of student success. These data may be used to create a typology of online students and of online course offerings. They may also be used to document student participation metrics, which may then be brought into longitudinal models as a kind of online student engagement measure. It may be useful to create categories of online students that document their level of participation and engagement on the web in a way that end-of-term data on grades and student ratings do not capture. This helps us understand of how successful institutions are at providing online learning and encourages the evaluation of programs and interventions.

Due to the proprietary and sometimes hidden nature of LMS data, as well as concerns about faculty control of the instructional process, access to LMS data is sometimes very limited. Reports about course-level, aggregate LMS data are more readily obtained, including data elements such as number of threads and postings and hours spent logged in by user. These data can then be merged with student and course data. Longitudinal student success models can certainly be used to understand this growing mode of delivery. It is not clear whether LMS data on MOOCs are available in the same type of reports and table structures for analysis.

## Section 10. Planning Issues

An ongoing program of longitudinal studies is needed with a clear agenda for "which studies to undertake at what time" and "how to present the results to decision-makers," explains Ewell (1987, p. 17). While there may be grand plans for a study, "A partially estimated model is better than none" (p. 17). These models can "become very complex," such as in the refinement of transfer data, and need to evolve (Lillibridge, 2008).

"It takes time to create…, but the results it produces will be well worth the effort" (Lillibridge, 2008, p. 30). Since longitudinal studies "require a good deal of time, effort, and other resources," "Always be prepared to commit enough of these scarce commodities" (Endo, 1992, p. 36). "One must allocate sufficient time and resources to resolve technical problems that can arise along the way and threaten the viability of implementation" (Pai et al, 2008, p. 21).

Musoba et al (2008, p. 108) point out that "many state agencies that collect student unit record data are too understaffed to make use of the data beyond basic reporting." Alliances such as the Wabash Study and the Indiana Project on Academic Success help make this kind of "extra-institutional research" possible. The Clearinghouse Signature Reports illuminate the potential gold mine of state and national-level SUR analyses. The upcoming release of SLDS reports brings this to life in the context of state-specific policies, economic conditions, and demographics.

The process of managing and presenting longitudinal tracking studies is documented by AIR (2010) and framed as questions to spend time on at the beginning, in the middle, at the end, and throughout a study. At the beginning, questions need to be "clear and understood by all stakeholders" and researchers need to ensure "that appropriate data are incorporated" (p. 16). Data definitions need to be consistent and documented. The schedule needs adequate time "to prepare datasets, including accessing, compiling, and cleaning them" (p. 16). There also need to be enough records or observations to conduct the study.

At the midpoint, it is important to anticipate surprises, such as finding that data are unavailable or in unusable formats. "Students sometimes behave in ways that defy logic." Also, "Stakeholders might develop sudden interest in variables that were not identified at the outset of the study" (AIR, 2010, p. 17). It must be understood that "a single study – even several studies – are unlikely to provide definitive answers to questions about student success" (p. 17).

When reaching the end of a study, the presentation of findings must be carefully considered. "Don't be seduced by fancy charts" and keep things "simple and to the point." The results need to "tell the story to decision-makers or help them to use data to make decisions" (AIR, 2010, p. 17).

Throughout the study, it is important to maintain and nurture relationships with data managers and IT staff. It is necessary to "keep focused on questions that matter, not those driven by idle curiosity or passing interest" (AIR, 2010, p. 18). Know that "You'll rarely find a 'right' way to answer to questions explored through longitudinal studies," so "do not let the quest for the perfect drive out the good" (p. 18).

As part of "changing the culture around data use," the DQC recommends that we "Build the capacity of all stakeholders to use longitudinal data for effective decision-making" (DQC, 2009, p. 7). This involves developing and implementing policies and practices for professional development to teach educators about how to access, analyze, and use longitudinal data appropriately. "Making the data available to educators is not enough to drive data use," however (p. 14). Professional development needs to take place "in multiple formats in a variety of venues." Incentives are needed to encourage participation and strategies that "raise awareness of available data and ensure that all key stakeholders, including state policymakers, know how to access, analyze and use the information" (p. 15).

Other obstacles and pitfalls to building a commitment to use data in this way are documented by ATD:

- Reporting requirements weigh heavily on colleges
- IR and IT functions are not widely visible or understood on many campuses
- Silos between IR and IT and between academic departments complicate coordination
- Making data accessible to a broad audience is difficult
- Concerns about data integrity inhibit widespread use
- Hiring people who don't share student success goals or have good communication skills
- Treating data as if it [sic]speaks for itself
- Making it impossible for the data to speak at all
- Hiding or ignoring "bad-news" data
- Using data as a hammer (ATD, 2012b, p. 6).

The right "data architect" is needed to help build an LDS and this may or may not be found within existing IT and IR staff. This role is separate from those involved in the business, information, applications, and technology architectures that must be addressed. The data architect must thoroughly master the complex data management and administration tasks that are usually done by others, leveraging this knowledge with new types of ETL and a vision for new data structures through a data warehouse or other forms of storage and retrieval that will serve longitudinal reporting and display. Fortunately, there is a great amount of attention being given now to this emerging role for business intelligence and analytics. Institutional researchers, while perhaps filling this role willingly or by default, must embrace this change and take the initiative and risk necessary to help lead LDS efforts in this direction.

This monograph is just a starting point, hopefully adding some context and references that will make this process easier and so that work is not done in a vacuum. This is an exciting time to be involved with these longitudinal projects and statewide efforts. There has never been more interest or a more highly developed set of tools to work with. The applications and future for longitudinal data analyses, if done well, are infinite and rewarding in a way that previous IR work had hoped for and that helped spur the profession to its unique role in the higher education community. Good luck!

# APPENDICES

## GLOSSARY OF ABBREVIATIONS

| Acronym | Title | Type |
|---|---|---|
| AACC | American Association of Community Colleges | National higher education association |
| AASCU | American Association of State Colleges and Universities | National higher education association |
| ABE | Adult Basic Education | Education level |
| ACT | ACT, Inc. | Non-profit organization for research and testing |
| AIR | Association for Institutional Research | National association |
| ANSWERS | Accessing National Surveys with Electronic Research Sources | Federal program/initiative, NCES |
| AP | Advanced Placement | Assessment/testing instrument/survey |
| APLU | Association of Public and Land-grant Universities | National higher education association |
| ARRA | American Recovery and Reinvestment Act | Federal funding stream |
| ATD | Achieving the Dream | Non-profit consortium |
| B&B | Baccalaureate and Beyond | Federal longitudinal study |
| BGSSE | Beginning College Survey of Student Engagement | Assessment/testing instrument/survey |
| BI | Business Intelligence | Information technology term |
| BLS | Bureau of Labor Statistics | Federal agency |
| BPS | Beginning Postsecondary Students | Federal longitudinal study |
| CAAP | Collegiate Assessment of Academic Proficiency | Assessment/testing instrument/survey |
| CAE | Council for Aid to Education | |
| CAL-PASS | California Partnership for Achieving Student Success | State agency |
| CASPAR | Computer-Aided Science Policy Analysis and | Federal software for data analysis, NSF |
| CCA | Complete College America | Non-federal data initiative |
| CCCR | | |
| CCSSE | Community College Survey of Student Engagement | Assessment/testing instrument/survey |
| CD | | |
| CEDS | Common Education Data Standards | Data quality initiative to standardize data |
| CEU | | |
| CHESS | Consortium for Higher Education Software Services | Data quality initiative to standardize data |
| CIP | | |
| CIRP | Cooperative Institutional Research Program | Assessment/testing instrument/survey |
| CLA | Collegiate Learning Assessment | Assessment/testing instrument/survey |

| Acronym | Title | Type |
|---|---|---|
| CMSS | Committee on Measures of Student Success | Federal program/initiative, ED |
| COMPETES | Creating Opportunities to Meaningfully Promote Excellence in Technology Education and Science Act | Federal funding stream |
| CPI | Consumer Price Index | Measure of inflation |
| CRC | | |
| CSEQ | College Student Experiences Questionnaire | Assessment/testing instrument/survey |
| CSRDE | Consortium for Student Retention Data | |
| CTE | Career and Technical Education | Taxonomy, disciplines |
| CUNY | | |
| DAS | Data Analysis System | Federal software for data analysis, NCES |
| DE | | |
| DQC | Data Quality Campaign | Data quality initiative to standardize data |
| DTS | | |
| ED | U.S. Department of Education | Federal department |
| EDRC | | |
| EEO | Equal Employment Opportunity | Federal program/initiative, taxonomy of occupations |
| EF | | |
| ELS:2002 | Education Longitudinal Study of 2002 | Federal longitudinal study |
| ESL | | |
| ESRA | Education Sciences Reform Act | Federal legislation |
| ETAA | | |
| ETL | Extract, Transform, and Load | IT term - data process |
| ETS | | |
| FAFSA | Free Application for Federal Student Aid | Federal data collection by ED FSA |
| FFEL | | |
| FERPA | Family Educational Rights and Privacy Act | Federal legislation |
| FSA | | |
| FTE | Full-Time Equivalent | Measure - student and personnel full-time equivalency |
| GE | | |
| GED | General Education Diploma | Education award |
| GIS | | |
| GPA | Grade Point Average | Student/course data element |
| GRS | | |
| HEGIS | Higher Education General Information System | Federal data collection by NCES, taxonomy of disciplines |
| HEOA | | |
| HERI | UCLA Higher Education Research Institute | Research center/institute |
| HHS | | |
| HLM | Higher Level Modeling | Statistical technique |

| Acronym | Title | Type |
|---|---|---|
| HS&B | High School and Beyond | Federal longitudinal study |
| HSLS:09 | High School Longitudinal Study of 2009 | Federal longitudinal study |
| ICLM | | |
| IDEA | Individuals with Disability Education Act | Federal legislation |
| IE | | |
| IES | Institute for Education Sciences | Federal agency |
| IHELP | | |
| ILD | Intensive Longitudinal Data | Longitudinal term for many measures over time |
| ILDS | | |
| IPAS | Indiana Project on Academic Success | Research center/institute |
| IPEDS | | |
| IR | Institutional Research | Higher education term |
| IT | | |
| JCAR | Joint Committee on Administrative Rules | State agency |
| JFF | | |
| K-12 | Kindergarten to High School | Education levels |
| KDD | | |
| LDAP | Lightweight Directory Access Protocol | IT term for security and directory standard |
| LDS | | |
| LMS | Learning Management System | Software for online learning |
| MAPP | | |
| MDM | Master Data Management | IT term for data administration approach |
| MOU | | |
| NAICS | North American Industry Classification System | Taxonomy, industries |
| NAICU | | |
| NASA | National Aeronautics and Space Administration | Federal department |
| NCES | | |
| NCHEMS | National Center for Higher Education Management Systems | Research center/institute |
| NCPR | | |
| NCSC | National Crosswalk Service Center | Federal program/initiative, ED |
| NDIR | | |
| NELS-88 | National Education Longitudinal Study of 1988 | Federal longitudinal study |
| NFES | | |
| NGA | National Governors Association | National association |

| Acronym | Title | Type |
|---|---|---|
| NLS-72 | National Longitudinal Study of the High | Federal longitudinal study |
| | National Postsecondary Education Cooperative | Federal program/initiative, NCES |
| NPSAS | | |
| NSC | National Student Clearinghouse | Research center/institute |
| NSF | | |
| NSLDS | National Student Loan Data System | Federal agency |
| NSSE | | |
| OBIEE | Oracle Business Intelligence Enterprise Edition | IT term for business intelligence software |
| ODBC | | |
| OLAP | On-line Analytical Processing | IT term for data processing and display |
| OMB | | |
| P-12 | Pre-Kindergarten to High School | Education levels |
| P-20 | | |
| Perkins IV | Carl D. Perkins Career and Technical Education Act of 2006 | Federal legislation |
| PESC | | Federal program/initiative, ED |
| PETC | Postsecondary Education Transcript Collections | Federal longitudinal study |
| PETS | Postsecondary Education Transcript Study | Federal longitudinal study |
| RCG | Recent College Graduates Survey | Federal longitudinal study |
| RHE | Research in Higher Education | Journal |
| RTTP | Race to the Top | Federal program/initiative, ED |
| S&P | Success and Progress Rate | Performance measure |
| SAS | Statistical Analysis System software | Software for data analysis |
| SAT | The College Board | Assessment/testing instrument/survey |
| SCHEV | State Council of Higher Education for Virginia | State agency |
| SDR | Survey of Doctoral Recipients | Federal longitudinal study |
| SED | Survey of Earned Doctorates | Federal data collection by NSF |
| SENSE | Survey of Entering New Student Engagement | Assessment/testing instrument/survey |
| SESTAT | Scientists and Engineers Statistical Data System | Federal software for data analysis, NSF |
| SFSF | State Fiscal Stabilization Fund | Federal funding stream |
| SHEEO | State Higher Education Executive Officers Association | National association |
| SIS | Student Information System | IT term for transaction-based data architecture |
| SLDS | Student Unit Record | IT term for individually identifiable student data |
| SLDS | State Longitudinal Data System | IT term for state unit record data system across the continuum |
| SMART | Science, Mathematics, and Research for Transformation | Taxonomy, disciplines |

| Acronym | Title | Type |
|---|---|---|
| SOC | Standard Occupational Category | Taxonomy, occupations |
| SQL | Structured Query Language | Programming syntax |
| SRK | | |
| SSN | Social Security Number | Personal identifier |
| STEM | | |
| SUNY | State University of New York | State, public system of higher education |
| SURE | | |
| TANF | Temporary Assistance for Needy Families | Federal assistance program |
| TEAL | | |
| TRP | Technical Review Panel | Federal program/initiative, NCES |
| TVEM | | |
| U-CAN | University and College Accountability Network | Non-federal data initiative |
| UI | | |
| VCCS | Virginia Community College System | State, public system of higher education |
| VFA | | |
| VITAL | Virginia Improves Teaching and Learning | State SLDS initiative for teacher data warehouse |
| VSA | | |
| WDQI | Workforce Data Quality Initiative | Federal program/initiative |
| WebCASPAR | | |
| WIA | Workforce Investment Act | Federal legislation |
| WIC | Special Supplemental Nutrition Program for Women, Infants and Children | Federal assistance program |
| WICHE | Western Interstate Commission for Higher Education | Non-profit consortium |
| WSBCTC | | |
| WSES | Wabash Student Experience Survey | Assessment/testing instrument/survey |
| WYSIWYG | What You See Is What You Get | IT term for visual design |

# References

AACC. (2011). Voluntary Framework of Accountability: Metrics Manual Version 1.0. Washington, D.C.: American Association of Community Colleges. Available online at: http://www.aacc.nche.edu/vfa.

AASCU. (2006). "Graduation Rates and Student Success: Squaring Means and Ends." *Perspectives* Fall 2006. Washington, D.C.: American Association of State Colleges and Universities.

Achieve. (2012). "Race to the Top: Accelerating College and Career Readiness in States." Achieve, Inc.

Adelman, Clifford. (1999). Answers in the tool box: Academic intensity, attendance patterns, and bachelor's degree attainment. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Adelman, Clifford. (2005). Moving into town—and moving on. The community college in the lives of traditional-age students. Washington, D.C.: Office of Vocational and Adult Education, U.S. Department of Education.

Adelman, Clifford. (2006). The toolbox revisited: Paths to degree completion from high school through college. Washington, DC: U.S. Department of Education.

AIR. (2012). "Common Education Data Standards." AIR Alert #39. Tallahassee, FL: Association for Institutional Research. Available online at: http://www.airweb.org/EducationAndEvents/IPEDSTraining/AdditionalResources/Pages/CEDS.aspx.

AIR. (2010). "Longitudinal Tracking for Institutional Research." Data & Decisions Academy. Tallahassee, FL: Association for Institutional Research.

Alliance. (2012). Committing to Quality: Guidelines for Assessment and Accountability in Higher Education. Washington, D.C.: New Leadership Alliance for Student Learning and Accountability.

Astin, Alexander W. (1997). "How 'Good' Is Your Institution's Retention Rate?" *Research in Higher Education*, 38(6). pp. 647-658.

Astin, Alexander W. and Lee, Jenny J. (2003). "How Risky Are One-Shot Cross-Sectional Assessments of Undergraduate Students?" *Research in Higher Education*, 44(6). pp. 657-672.

ATD. (2012a). "Assessment of Evidence Series." Achieving the Dream. Available online at: http://www.achievingthedream.org/resources/ccrc_assessment_evidence.

ATD. (2012b). Building Institutional Capacity for Data-Informed Decision Making." Achieving the Dream, Cutting Edge Series, No. 3.

Bahr, Peter R. (2010). "The Bird's Eye View of Community Colleges: A Behavioral Typology of First-Time Students Based on Cluster Analytic Classification. *Research in Higher Education*, 51. pp. 724-749.

Bahr, Peter R. et al. (2011a). "A Typology of Students' Use of the Community College." In *New Directions for Institutional Research*, Assessment Supplement 2011. pp. 33-48.

Bahr, Peter R. et al. (2011b). "The Use of Cluster Analysis in Typological Research on Community College Students." In *New Directions for Institutional Research*, Assessment Supplement 2011. pp. 67-81.

Bauer, Karen W. (2004). "Conducting Longitudinal Studies." In *New Directions for Institutional Research*, spring 2004, no. 121. Overcoming Survey Research Problems. pp. 75-.

Bergeron, David. (2012a). "Gainful Employment Electronic Announcement #42 - 2011-2012 Disclosures for Gainful Employment Programs. Washington, D.C.: U.S. Department of Education. Available online at: http://ifap.ed.gov/eannouncements/112312GEAnnounce42DisclosureofGEPrograms1112.html.

Bergeron, David. (2012b). "Gainful Employment and the White House College Scorecard." Presentation at the at the 2012 Annual Forum of the Association for Institutional Research, New Orleans, LA.

Berkner, Lutz. (2000). "Using National Education Longitudinal Study Data to Examine the Transition to College." In *New Directions for Institutional Research*, Fall 2000, no. 107. Understanding the College Choice of Disadvantaged Students. pp. 103-107.

Borden, Victor M. H. et al. (1996). Setting a census date to optimize enrollment, retention, and tuition revenue projects. AIR Professional File, no. 62. Tallahassee, FL: Association for Institutional Research.

Boughan, Karl. (2000). "The Role of Academic Process in Student Achievement: An Application of Structural Equations Modeling and Cluster Analysis to Community College Longitudinal Data." AIR Professional File, no. 74. Tallahassee, FL: Association for Institutional Research.

Bowman, Nicholas A. and Small, Jenny L. "Do College Students Who Identify with a Privileged Religion Experience Greater Spiritual Development? Exploring Individual and Institutional Dynamics." *Research in Higher Education*, 51. pp. 595-614.

Brown, Keith. (2009). "A Better Way to Measure Community College Performance: An Achieving the Dream Cross-State Data Initiative." Presentation at the 2009 State Higher Education Officers/National Center for Education Statistics Network Conference and IPEDS Workshop.

Bryant, Alyssa N. (2011). "Ecumenical Worldview Development by Gender, Race, and Worldview: A Multiple-Group Analysis of Model Invariance." *Research in Higher Education,* 52. pp. 460-479.

Buckley, Jack. (2012). "What You Need to Know about CEDS and Version 2.0." Panel presentation at the Data Qality Meeting How the New Common Education Data Standards Can Support Policy and Practice, February 23, 2012. Available online at:
http://www.dataqualitycampaign.org/files/events/resources/DQC%20CEDS%20Webinar%202_23_12.pdf.

CCRC. (2012b). "Research Overview." New York: Community College Research Center, Teachers College, Columbia University. Available online at: http://ccrc.tc.columbia.edu/Research.asp.

CEDS. (2012). "CEDS Data Model – Introduction." U.S. Department of Education, National Center for Education Statistics, Common Education Data Standards. Available online at: http://ceds.ed.gov/dataModel.aspx.

Chen, Pu-Shih D. and Cragg, Kristina. (2012). "Multilevel Modeling: Applications to Research on the Assessment of Student Learning, Engagement, and Developmental Outcomes." In *New Directions for Institutional Research*, no. 154. Multilevel Modeling Techniques and Applications. pp. 95-109.

Clagett, Craig. (1992). "Enrollment Management." In Whitely, Meredith A., Porter, John D., and Fenske, Robert H. (eds.), The Primer for Institutional Research. Tallahassee, FL: Association for Institutional Research. pp. 12-24.

CMSS. (2011). "A Report to Secretary of Education Arne Duncan." U.S. Department of Education, Committee on Measures of Student Success.

Colcagno, Juan Carolos et al. (2007). "Stepping Stones to a Degree: The Impact of Enrollment Pathways and Milestones on Community College Student Outcomes." *Research in Higher Education*, 48(7). pp. 775-801.

Cole, James S. and Korkmaz, Ali. (2010). "Using Longitudinal Data to Improve the Experiences and Engagement of First-Year Students." In *New Directions for Institutional Research*, Assessment Supplement 2010. pp. 43-51.

Completion Arch. (2012). "The Completion Arch: Measuring Community College Student Success." College Board. Available online at: http://completionarch.collegeboard.org.

Conklin, Karen A. (1995). "Community College Students' Persistence and Goal Attainment: A 5-Year Longitudinal study." AIR Professional File, no. 55. Tallahassee, FL: Association for Institutional Research.

CSC. (2012). "State Authorization Reciprocity Agreement." April, 2012 – Working Draft. Lexington, KY: Council of State Governments, The National Center for Interstate Compacts.

Cunningham, Alicia and Milam, John. (2005). Feasibility of a Student Unit Record System within the Integrated Postsecondary Data System. NCES 2005–160. U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Deng, Heping. (2006). "Are There Differences between Transfers from Community College Career-Oriented Programs and Liberal Arts Programs?" IR Applications, vol. 11. Tallahassee, FL: Association for Institutional Research.

DesJardins, Stephen L. (2003). "Event History Methods: Conceptual Issues and An Application to Student Departure from College."‖ *Higher Education: Handbook of Theory and Research XVIII*: 421-471.

Diaz, Alicia A. (2010). "SREB States Sharing Data: Policies that Work." Presentation at the 2010 State Higher Education Officers/National Center for Education Statistics Network Conference and IPEDS Workshop.

DQC. (2009a). "DQC 2009 Annual Survey Update and State Progress Report." Washington, D.C.: Data Quality Campaign.

DQC. (2009b). "The Next Step: Using Longitudinal Data Systems to Improve Student Success." Washington, D.C.: Data Quality Campaign.

DQC. (2010). "Education and Workforce Data Connections: A Primer on States' Status." Washington, D.C.: Data Quality Campaign.

DQC. (2011). "States Could Empower Stakeholders to Make Education Decisions with Data… but They Haven't Yet." Washington, D.C.: Data Quality Campaign, *Data for Action 2011*.

DQC. (2011). "Just the Facts: State of the States' Efforts to Support Effective Data Use." Washington, D.C.: Data Quality Campaign, *Data for Action 2011*.

DQC. (2012c). "Roadmap." Washington, D.C.: Data Quality Campaign. Available online at: http://www.DataQualityCampaign.org/Roadmap.

DQC. (2012a). "State Fiscal Stabilization Fund – Revised Data Collection and Reporting Requirements: DQC Summary." Washington, D.C.: Data Quality Campaign.

DQC. (2012b). "Leveraging Federal Funding for Longitudinal Data Systems: A Roadmap for States." Washington, D.C.: Data Quality Campaign. Available online at: http://www.dataqualitycampaign.org/build/fedfunding.

ED. (2012). "Action Plan for Improving Measures of Postsecondary Student Success." Washington, D.C.: U.S. Department of Education.

ED. (2009). "Statewide Longitudinal Data Systems." Washington, D.C.: U.S. Department of Education. Available online at: http://www2.ed.gov/programs/slds/factsheet.html.

Elfman, Lois. (2012). "U.S. Department of Education Develops College Scorecard." *Diverse Issues in Higher Education*, April 18, 2012.

Endo, Jean. (1992). "Student Impacts." In Whitely, Meredith A., Porter, John D., and Fenske, Robert H. (eds.), The Primer for Institutional Research. Tallahassee, FL: Association for Institutional Research. pp. 25-36.

Engberg, Mark E. and Wolniak, Gregory C. (2010). "Examining the Effects of High School Contexts on Postsecondary Enrollment." *Research in Higher Education*, 51. pp. 132-153.

Ewell, Peter. (2008). Community College Data and Performance Measure Toolkit: Information, Strategies and Resources for State Policymakers and Community College Leaders. Austin, TX: Community College Bridges to Opportunity Initiative, University of Texas at Austin.

Ewell, Peter. (1987). "Principles of Longitudinal Enrollment Analysis: Conducting Retention and Student Flow studies." In Muffo, John A. and McLaughlin, Gerald W. (eds.), A Primer on Institutional Research. Tallahassee, FL: Association for Institutional Research. pp. 1-19.

Ewell, Peter and Boeke, M. (2007). *Critical Connections: Linking States' Unit Record Systems to Track Student Progress*. Indianapolis, IN: Lumina Foundation for Education.

Ewell, Peter and Jenkins, David. (2008). Using state student unit record data to increase community college student success.‖ *New Directions for Community Colleges,* #143, pp. 71-81.

Ewell, Peter and L'Orange, Hans. (2009). "The Ideal State Postsecondary Data System: 15 Essential Characteristics and Required Functionality." Boulder, CO: National Center for Higher Education Management Systems and State Higher Education Executive Officers.

Ewell, Peter et al. (1988). *Establishing a Longitudinal Student Tracking System: An Implementation Handbook*. Boulder, CO: National Center for Higher Education Management Systems.

Ewell, Peter et al. (2003). *Following the Mobile Student: Can We Develop the Capacity for a Comprehensive Database to Assess Student Progression?* Indianapolis: Lumina Foundation for Education.

Fenske, Robert H. et al. (2000). "Tracking Financial Aid and Persistence of Women, Minority, and Needy Students in Science, Engineering, and Mathematics." *Research in Higher Education*, 41(1). pp. 67-94.

Fike, David S. and Fike, Renea. (2008) "Predictors of First-Year Student Retention in the Community College." *Community College Review*, 26(2). pp. 68-88.

GAO. (2009). "Career and Technical Education: States Have Broad Flexibility in Implementing Perkins IV." GAO Highlights, GAO-09-683. Washington, D.C.: U.S. Government Accountability Office.

Glover, Robert H. and Wilcox, Jerry. (1992). "An Interactive Model for Studying Student Retention." AIR Professional File, no. 44. Tallahassee, FL: Association for Institutional Research.

Gould, Tate and Anthony, Emily. (2010). "SLDS Statewide Longitudinal Data Systems Grant Programs." PowerPoint Presentation. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics. Available online at: http://nces.ed.gov/programs/SLDS.

Guerin, Keith. (2009). "Constructing, Managing and Analyzing Longitudinal Data: Programming Efficiencies with SAS Software." Pre-Forum Workshop presentation at the 2008 Annual Forum of the Association for Institutional Research, Atlanta, GA.

Hagerty, John et al. (2012). "Magic Quadrant for Business Intelligence Platforms." Gartner. Research G00225500.

Herzog, Serge. (2011). "Gauging Academic Growth of Bachelor Degree Recipients: Longitudinal vs. Self-Reported Gains in General Education." In *New Directions for Institutional Research*, no. 150. Validity and Limitations of College student Self-Report Data. pp. 21-39.

Hossler, Don et al. (2012a). "Transfer and Mobility: A National View of Pre-Degree Student Movement in Postsecondary Institutions." Signature Report, no. 2. Herndon, VA: National Student Clearinghouse Research Center and the Project on Academic Success, Indiana University.

Hossler, Don et al. (2012b). "Reverse Transfer: A National View of Student Mobility from Four-Year to Two-Year Institutions." Signature Report, no. 3. Herndon, VA: National Student Clearinghouse Research Center and the Project on Academic Success, Indiana University.

Howson, Cindi. (2012). BI Scorecard. Available online at http://www.biscorecard.com/.

Hu, Shouping and Li, Shaoqing. (2011). "Typological Research on College Students for Better Outcomes." In *New Directions for Institutional Research*, Assessment Supplement 2011. pp. 83-87.

Hu, Shouping and Wolniak, Gregory C. (2010). "Initial Evidence on the Influence of College Student Engagement on Early Career Earnings." *Research in Higher Education*, 51. pp. 750-766.

Hu, Shouping et al. (2011). "Student Typologies in Higher Education." In *New Directions for Institutional Research*, Assessment Supplement 2011. pp. 5-15.

Hughes, Katherine L. et al. (2012). "Broadening the Benefits of Dual Enrollment: Reaching Underachieving and Underrepresented Students with Career-Focused Programs." New York: Community College Research Center, Teachers College, Columbia University.

Hurtado, Sylvia et al. (2008). "Training Future Scientists: Predicting First-Year Minority Student Participation in Health Science Research." *Research in Higher Education*, 49. pp. 126-152.

IPEDS TRP. (2008). "Report and Suggestions from IPEDS Technical Review Panel #24: Collecting GRS Data for Part-Time Students and Pell Grant Recipients." Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

IPEDS TRP. (2012). "Report and Suggestions from IPEDS Technical Review Panel #37: Selected Outcomes of the Advisory Committee on Measures of Student Success." Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Ishitani, Terry T. (2008). "How to Explore Timing of Intervention for Students at Risk of Departure." In *New Directions for Institutional Research*, spring 2008, no. 137. Alternative Perspectives in Institutional Planning. pp. 105-122.

Ishitani, Terry. (2003). "A Longitudinal Approach to Assessing Attrition Behavior Among First-Generation Students: Time-Varying Effects of Pre-College Characteristics. *Research in Higher Education*, 44(4). pp. 433-449.

Ishitani, Terry. (2008). "How Do Transfer Survive after 'Transfer Shock'? A Longitudinal Study of Transfer Student Departure at a Four-Year Institution." *Research in Higher Education*, 49. pp. 403-419.

Ishitani, Terry (2006). "Longitudinal Effects of College Preparation Programs on College Retention." IR Applications, vol. 9. Tallahassee, FL: Association for Institutional Research.

Ishitani, Terry and DesJardins, Stephen L. (2002). "A Longitudinal Investigation of Dropout from College in the United States."‖ *Journal of College Student Retention*, 4(2): 173-201.

JCAR. (1996). Joint Commission on Accountability Reporting, Technical Conventions Manual. Joint Commission on Accountability Reporting, American Association of State Colleges and Universities, American Association of Community Colleges, and National Association of State Universities and Land Grant Colleges. Washington, D.C.: American Association of State Colleges and Universities.

Jenkins, Davis. (2008). "A Short Guide to 'Tipping Point' Analyses of Community College Student Labor Market Outcomes." *CCRC Research Tools*, no. 3, April 2008. New York: Community College Research Center, Teachers College, Columbia University.

Jenner, Carol B. (2009). "Washington's Education Research and Data Center." Presentation at the 2009 State Higher Education Officers/National Center for Education Statistics Network Conference and IPEDS Workshop.

JFF. (2010). "The Developmental Education Initiative: State Policy Framework and Strategy." Developmental Education Initiative. Boston: Jobs for the Future.

JFF. (2009). "Setting Up Success in Developmental Education: How State Policy Can Help Community Colleges Improve." Developmental Education Initiative. Boston: Jobs for the Future.

JFF. (2008). "Test Drive: Six States Pilot Better Ways to Measure and Compare Community College Performance. An Achieve the Dream Policy Brief." Boston: Jobs for the Future.

Jones-White, Daniel R. et al. (2008). "Redefining Student Success: Assessing Different Multinomial Regression Techniques for the Study of Student Retention and Graduation Across Institutions of Higher Education." Presentation at the 2008 Annual Forum of the Association for Institutional Research, Seattle, WA.

Karp, Melinda M. et al. (2012). "College 101 Courses for Applied Learning and Student Success." CCRC Working Paper No. 49. New York: Community College Research Center, Teachers College, Columbia University.

Keller, Christine, and Gore, Paul. (2012). "The Future of Student Learning Outcomes Assessment within the VSA." Presentation at the annual conference of the Rocky Mountain Association for Institutional Research, Laramie, WY.

Keup, Jennifer R. (2004). "The Cooperative Institutional Research Program Freshman Survey and Your First College Year: Using Longitudinal Data to Assess the First Year of College." *Assessment Update*, 16(2):2-11.

King, Joe P. et al. (2012). "Using Existing Data Sources/Programs and Multilevel Modeling Techniques for Questions in Institutional Research." In *New Directions for Institutional Research*, no. 154. Multilevel Modeling Techniques and Applications. pp. 41-55.

Knight, William E. (2004). "Time to Bachelor's Degree Attainment: An Application of Descriptive, Bivariate, and Multiple Regression Techniques." IR Applications, vol. 2. Tallahassee, FL: Association for Institutional Research.

Knight, William E. (2002). "Toward a Comprehensive Model of Influences Upon Time to Bachelor's Degree Attainment." AIR Professional File, no. 85. Tallahassee, FL: Association for Institutional Research.

Leinbach, D. Timothy and Jenkins, Davis. (2008). "Using Longitudinal Data to Increase Community College Student Success: A Guide to Measuring Milestone and Momentum Attainment." *CCRC Research Tools*, no. 2, Spring 2008. New York: Community College Research Center, Teachers College, Columbia University.

Lesik, Sally A. (2007). "Do Developmental Mathematics Programs Have A Causal Impact on Student Retention? An Application of Discrete-Time Survival and Regression-Discontinuity Analysis." *Research in Higher Education*, 48(5). pp. 583-608.

Lietz, Petra and Mathews, Bobbie. (2010). "The Effects of College Students' Personal Values on Changes in Learning Approaches." *Research in Higher Education*, 51. pp. 65-87.

Lillibridge, Fred. (2008). "Retention Tracking Using Institutional Data." In *New Directions for Institutional Research*, fall 2008, no. 143. Student Tracking in the Community College. pp. 19-30.

Long, Bridget Terry and Kurlaender, Michal. (2008). "Do Community Colleges Provide a Viable Pathway to a Baccalaureate Degree?" NBER Working Paper #14367. Cambridge, MA: National Bureau of Economic Research.

L'Orange, Hans P. (2012). "The Intersection of Public Policy and Data in Higher Education." Presentation at the 2012 conference of the Rocky Mountain Association for Institutional Research, Laramie, WY.

L'Orange, Hans P. et al. (2010). "State Data Systems: A Resource for Policy Development." Panel presentation at the 2010 Annual Forum of the Association for Institutional Research, Chicago, IL.

Luan, Jing et al. (2012). "Exploring and Mining Data." In Howard, Richard D., McLaughlin, Gerald W., and Knight, William E. (eds)., The Handbook of Institutional Research. San Francisco, CA: Jossey-Bass. pp. 478-501.

Luan, Jing et al. (2009). "Using a Data Mining Approach to Develop a Student Engagement-Based Institutional Typology." IR Applications, vol. 18. Tallahassee, FL: Association for Institutional Research.

Lumina. (2012). A Stronger Nation through Higher Education: How and Why Americans Must Achieve a Big Goal for College Attainment. Indianapolis: Lumina Foundation for Education.

Martin, Georgianna I. and Loes, Chad N. (2010). "What Incentives Can Teach Us About Missing Data in Longitudinal Assessment." In *New Directions for Institutional Research*, Assessment Supplement 2010. pp. 17-28.

McLaughlin, Virginia L. et al. (2008). "Virginia Improves Teaching and Learning (VITAL): A Comprehensive Data System for Teacher Quality." In *New Directions for Institutional Research*, fall 2008, no. 139. Conducting Institutional Research in Non-Campus-Based Settings. pp. 83-93.

Menard, S. (1991). Longitudinal Research. Quantitative Applications in the Social Sciences, paper no. 07–076. Newbury Park, CA: Sage.

Microsoft. (2012a). "Power View (SSRS)." Available online at [http://technet.microsoft.com/en-us/library/hh213579(d=printer).aspx](http://technet.microsoft.com/en-us/library/hh213579(d=printer).aspx).

Microsoft. (2012b). "What's New in PowerPivot." Available online at http://technet.microsoft.com/en-us/library/hh272053(d=printer).aspx.

Microsoft. (2012c). "Getting Started with Report Builder." Available online at http://technet.microsoft.com/en-us/library/dd220460(d=printer).aspx.

Microsoft. (2012b). "What's New in PowerPivot." Available online at http://technet.microsoft.com/en-us/library/hh272053(d=printer).aspx.

MicroStrategy. (2012). "MicroStrategy Reporting Suite."

Middaugh, Michael F. (1992). "Persistence." In Whitely, Meredith A., Porter, John D., and Fenske, Robert H. (eds.), The Primer for Institutional Research. Tallahassee, FL: Association for Institutional Research. pp. 1-11.

Milam, John. (2009a). "New Measures of Student Success in Public Higher Education." Stephens City, VA: HigherEd.org, Inc.

Milam, John. (2008a). "Noncredit Instructional Activities and IPEDS." IPEDS Research and Development Background Paper #1. Prepared under the guidance of the National Postsecondary Education Cooperative for the IPEDS Technical Review Panel. Stephens City, VA: HigherEd.org, Inc.

Milam, John. (2008b). "Nontraditional Students in Public Higher Education: A Multi-State, Student Unit Record Study." Stephens City, VA: HigherEd.org, Inc.

Milam, John. (2009b). "Nontraditional Student Success in Public Two-year Colleges." Stephens City, VA: HigherEd.org, Inc.

Milam, John. (2006). Ontologies in higher education. In A. S. Metcalfe (Ed.), Knowledge management and higher education: A critical analysis (pp. 34–61). Hershey, PA: Information Science Publishing.

Milam, John. (2005). The role of noncredit courses in serving non-traditional learners.‖ In Pusser, B. (Ed.), *Arenas of entrepreneurship: Where nonprofit and for-profit institutions compet*e. New Directions for Higher Education, 129. San Francisco: Jossey-Bass, pp. 55-68.

Milam, John et al. (2012). "Business Intelligence and Analytics: The IR Vision for Data Administration, Reporting, Data Marts, and Data Warehousing." In Howard, Richard D., McLaughlin, Gerald W., and Knight, William E. (eds)., The Handbook of Institutional Research. San Francisco, CA: Jossey-Bass. pp. 434-453.

Milam, John et al. (2008). *The Cost of Not Knowing: State Capacity for Postsecondary Cost Studies*. Stephens City, VA: HigherEd.org, Inc.

Milam, John. (2009c). "Using the VCCS SAS Data Warehouse." Presentation at the 2009 meeting of the VCCS IR Workgroup."

Miller, Adam. (2012). "Ed-Fi ." Panel presentation for February 23, 2012 Webinar, How the New Common Education Data Standards Can Support Policy and Practice." Washington, D.C.: Data Quality Campaign. Available online at: http://www.dataqualitycampaign.org/events/details/301.

Miller, Troy. (2012). "Study Shows Dual Enrollment Has Positive Influence on degree Completion for Low-Income Students." Florida College Access Network.

Mosley, Mark and Brackett, Michael, (Eds.). (2010). DAMA guide to the data management body of knowledge. Bradley Beach, NJ: Technics Publications.

Musoba, Glenda D. et al. (2008). "Confronting Ambiguity, Anarchy, and Crisis in Institutional Research: Using Student Unit Record Databases in Extra-Institutional Research." In *New Directions for Institutional Research*, Fall 2008, no. 139. Conducting Institutional Research in Non-Campus-Based Settings. pp. 95-116.

NCES. (2012). "Statewide Longitudinal data Systems Grant Program." U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Available online at: http://nces.ed.gov/Programs/SLDS/p20w.asp.

NCES. (2010a). "Basic Concepts and Definitions for Privacy and Confidentiality in Student Education Records." SLDS Technical Brief, no. 1. NCES 2011-601. Washington, D.C.: , Institute of Education Sciences, National Center for Education Statistics.

NCES. (2010b). "Data Stewardship: Managing Personally Identifiable Information in Electronic Student Education Records." SLDS Technical Brief, no. 2. NCES 2011-602. Washington, D.C.: Institute of Education Sciences, National Center for Education Statistics.

NCES/BLS. (2011). "Guidelines for Using the CIP-SOC Crosswalk." Washington, D.C.: National Center for Education Statistics and Bureau of Labor Statistics. Available online at: http://nces.ed.gov/ipeds/cipcode/resources.aspx?y=55.

NCHEMS. (2008). Tracking Postsecondary Students Across State Lines: Results of a Pilot Multi-State Data Exchange Initiative.‖ Boulder, CO: National Center for Higher Education Management Systems.

NCHEMS/SHEEO. (2008). "Focus on the Ideal State Postsecondary Data System."‖ *Network News* 27(1). Boulder, CO: State Higher Education Executive Officers, National Center for Higher Education Management Systems. Available online at: http://www.sheeo.org/network/netnews/nn-v27-n1.pdf.

NCSC. (2012b). "NSCSC Prototype Crosswalk files." Des Moines, IA: U.S. Department of Education, National Crosswalk Service Center. Available online at: http://xwalkcenter.org/.

NCSC. (2012b). "Crosswalk File Documentation." Des Moines, IA: U.S. Department of Education, National Crosswalk Service Center. Available online at: http://webdata.xwalkcenter.org/ftp/download/xwalks/occindo.pdf.

NFES. (2010a). Traveling Through Time: The Forum Guide to Longitudinal Data Systems. Book One of Four: What is an LDS? NFES 2010-805. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

NFES. (2010b). Traveling Through Time: The Forum Guide to Longitudinal Data Systems. Book Two of Four: Planning and Developing an LDS. NFES 2010-804. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

NFES. (2011a). Traveling Through Time: The Forum Guide to Longitudinal Data Systems. Book Three of Four: Effectively Managing LDS Data. NFES 2011-805. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

NFES. (2011b). Traveling Through Time: The Forum Guide to Longitudinal Data Systems. Book Four of Four: Advanced LDS Usage. NFES 2011-802. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Noble, Julie and Sawyer, Richard. (2012). "Institutional Research with Published Instruments." In Howard, Richard D., McLaughlin, Gerald W., and Knight, William E. (eds)., The Handbook of Institutional Research. San Francisco, CA: Jossey-Bass. pp. 523-549.

NSC. (2011a). "Mobility Snapshot Report." Herndon, VA: National Student Clearinghouse Research Center.

NSC. (2011b). "Persistence Snapshot Report." Herndon, VA: National Student Clearinghouse Research Center.

NSC. (2012). "What We Do." Herndon, VA: National Student Clearinghouse Research Center.

NSLDS. (2011). Gainful Employment User Guide. Version 1.0. Washington, D.C.: U.S. Department of Education, National Student Loan Data System.

Olsen, Michelle D. (2010). "Modeling a Longitudinal Relational Research Data System." Doctoral dissertation, University of North Dakota. UMI #3455230.

O'Connell, Ann A. and Reed, Sandra J. (2012). "Hierarchical Data Structures, Institutional Research, and Multilevel Modeling." In *New Directions for Institutional Research*, no. 154. Multilevel Modeling Techniques and Applications. pp. 5-22.

O'Toole, Dennis M. et al. (2003). "A Longitudinal Analysis of the Frequency of Part-time Enrollment and the Persistence of Students Who Enroll Part-time." *Research in Higher Education*, 44(5). pp. 519-537.

Padgett, Ryan D. et al. (2012). "First-Generation Undergraduate Students and the Impacts of the First Year of College: Additional Evidence." *Journal of College Student Development*, 53(2). pp. 243-266.

Padgett, Ryan D. et al. (2010). "Required, Practical, or Unnecessary? An Examination and Demonstration of Propensity Score Matching Using Longitudinal Secondary Data." In *New Directions for Institutional Research*, Assessment Supplement 2010.

Pai, Wendell et al. (2008). "Developing a Statewide Student Tracking Tool." In *New Directions for Institutional Research*, fall 2008, no. 139. Conducting Institutional Research in Non-Campus-Based Settings. pp. 5-25.

Pascarella, Ernest T. (2006). "How College Affects Students: Ten Directions for Future Research." *Journal of College Student Development* 45(5): pp. 508-520.

Pascarella, Ernest T. et al. (2010). "How Effective are the NSSE Benchmarks in Predicting Important Educational Outcomes?" *Change*, January-February 2010. Available online at: http://www.changemag.org/Archives/Back Issues/January-February 2010/full how-effective.html

Pascarella, Ernest T. and Terenzini, Patrick T. (2005). How College Affects Students: A Third Decade of Research (Vol 2). San Francisco, CA: Jossey-Bass.

Penn State. (2012). "Intensive Longitudinal Data (ILD). State College, PA: Pennsylvania State University, The Methodology Center. Available online at: http://methodology.psu.edu/ra/inten-longit-data.

Perry, Raymond P. et al. (2000). "Faculty in Transition: A Longitudinal Analysis of Perceived Control and Type of Institution in the Research Productivity of Newly Hired Faculty." *Research in Higher Education*, 41(2). pp. 165-194.

PESC. (2012). "PESC: Mission & History." Washington, D.C.: Postsecondary Electronic Standards Council. Available online at: http://www.pesc.interior.php?page_id=156.

Phillips, Jennifer. (2009). "Building and Strengthening State Data Systems to Measure Community College and Workforce Outcomes." Policy Brief, The Working Poor Families Project. Available online at: http://www.workingpoorfamiliaes.org.

Pike, Gary R. et al. (2011). "Using Instrumental Variables to Account for Selection Effects in Research on First-Year Programs." *Research in Higher Education*, 52. pp. 194-214.

Pinkus, Lyndsay. (2012). "How the New Common Education Data Standards Can Support Policy and practice." Panel presentation at the State Educational Technology Directors Association (SETDA) Leadership Conference.

Porter, Stephen R. (2002). "Including Transfer-Out Behavior in Retention Models: Using the NSC EnrollmentSearch Data." AIR Professional File, no. 82. Tallahassee, FL: Association for Institutional Research.

Prescott, Brian T. (2012). "Beyond Borders: Linking Data Systems Across Sectors and States." Presentation at the 2012 Annual Forum of the Association for Institutional Research in New Orleans, LA.

Prescott, Brian T. and Bensimon, Estella M. (2011). "We have goals… now what? Using Data to Achieve Equity in Educational Outcomes." Presentation at the 2011 State Higher Education Officers/National Center for Education Statistics Network Conference and IPEDS Workshop.

Prescott, Brian T. and Ewell, Peter. (2009). "A Framework for a Multi-State Human Capital Development Data System." Boulder, CO: Western Interstate Commission for Higher Education.

Price, Derek V. (2004). "Educational Debt Burden Among Student Borrowers: An Analysis of the Baccalaureate & Beyond Panel, 1997 Follow-Up." *Research in Higher Education*, 45(7). pp. 701-737.

Priest, Elisa L. and Collinsworth, Ashley W. (2011). "Intro to Longitudinal Data: A Grad Student 'How-To Paper." Grad Student How-To Series. South Central SAS Users Group.

Prince, David and Jenkins, Davis. (2005). "Building Pathways to Success for Low-Skill Adult Students: Lessons for Community College Policy and Practice from a Longitudinal Student Tracking Study." *CCRC Brief*, no. 25, April, 2005. New York: Community College Research Center, Teachers College, Columbia University.

Pusser, Brian et al. (2007). Returning to Learning: Adults' Success in College Is Key to America's Future. Lumina Foundation for Education New Agenda Series. Indianapolis, IN: Lumina Foundation.

Reindl, Travis. (2011). "Complete to Compete: The 2010-11 NGA Chair's Initiative. Presentation at the National Governors Association Center for Best Practices.

Reynolds, Pauline J. et al. (2010). "Using Longitudinal Mixed-Methods Research to Look at Undeclared Students." In *New Directions for Institutional Research*, Assessment Supplement 2010. pp. 53-66.

Rice, Gary A. and Russell, Alene B. (2012). "Refocusing Student Success: Toward a Comprehensive Model." In Howard, Richard D., McLaughlin, Gerald W., and Knight, William E. (eds), The Handbook of Institutional Research. San Francisco, CA: Jossey-Bass. pp. 237-255.

Rocconi, Louis M. and Ethington, Corinna A. (2009). "Assessing Longitudinal Change: Adjustment for Regression to the Mean Effects." *Research in Higher Education*, 50. pp. 368-376.

Rodriguez, Olga et al. (2012). "Bridging College and Careers: Using Dual Enrollment to Enhance Career and Technical Education Pathways." NCPR Working Paper, July 2012. New York: Community College Research Center, National Center for Postsecondary Research, Teachers College, Columbia University.

Romano, R. M., & Wisniewski, M. (2005). Tracking community college transfers using National Student Clearinghouse data. AIR Professional File, no. 94. Tallahassee, FL: Association for Institutional Research.

Ruddock, Maryann. (2012). "Developing K-20+ State Databases." In Howard, Richard D., McLaughlin, Gerald W., and Knight, William E. (eds)., The Handbook of Institutional Research. San Francisco, CA: Jossey-Bass. pp. 404-419.

Ruffing, Katherine. (2006). "The History of Career Clusters."‖ Washington, D.C.: National Association of State Directors of Career Technical Education Consortium.

Russell, Alene B. (1999). "The Status of Statewide Student Transition Data Systems: A Survey of SHEEO Agencies." Boulder: State Higher Education Executive Officers.

Schaff, Mark and Janis, Les. (2009). "Occupational Supply and Demand System (OSDS): Design Improvements and Recent Applications." Presentation at the 2009 State Higher Education Officers/National Center for Education Statistics Network Conference and IPEDS Workshop.

Schenk, Tom Jr. (2011). "Measuring Transitions into the Workforce as a Form of Accountability." IR Applications, vol. 32. Tallahassee, FL: Association for Institutional Research.

SCHEV. (2012). "Post-Completion Wages of Graduates." Richmond, VA: State Council of Higher Education for Virginia.

Seifert, Tricia A. (2010). "Editor's Notes." In *New Directions for Institutional Research*, Assessment Supplement 2010. pp. 1-3.

Seifert, Tricia A. et al. (2010). "The Importance of Longitudinal Pretest-Posttest Designs in Estimating College Impact." In *New Directions for Institutional Research*, Assessment Supplement 2010. pp. 3-16.

Shapiro, Doug et al. (2012). "Completing College: A National View of Student Attainment Rates." Signature Report, no. 4. Herndon, VA: National Student Clearinghouse Research Center and the Project on Academic Success, Indiana University.

Singer, Judith D. and Willett, John B. (2003). Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. New York City, NY: Oxford University Press.

Smart, John C. and Feldman, Kenneth A. (1998). "'Accentuation Effects' of Dissimilar Academic Departments: An Application and Exploration of Holland's Theory." *Research in Higher Education*, 39(4). pp. 385-418.

Sommo, Colleen et al. (2012). "Commencement Day: Six-Year Effects of a Freshman Learning Community Program at Kingsborough Community College." MDRC.

Stupnisky, Robert H. et al. (2008). "The Interrelation of First-Year College Students' Critical Thinking Disposition, Perceived Academic Control, and Academic Achievement." *Research in Higher Education*, 49. pp. 513-530.

TechNet Wiki. (2012). "PowerPivot Overview." TechNet wiki document. Available online at: http://social.technet.microsoft.com/wiki/contents/articles277.powerpivot-overview.aspx.

Terenzini, Patrick T. (2010). "Assessment with Open Eyes: Pitfalls in Studying Student Outcomes." In *New Directions for Institutional Research*, Assessment Supplement 2009. Assessing Student Outcomes: Why, Who, What, How?

Terenzini, Patrick T. (1987). "Studying Student Attrition and Retention." In Muffo, John A. and McLaughlin, Gerald W. (eds.), A Primer on Institutional Research. Tallahassee, FL: Association for Institutional Research. pp. 20-35.

Thomas, Charles R. (2004a). *CHESS data definitions for colleges and universities*. Boulder, CO: National Center for Higher Education Management Systems.

Thomas, Charles R. (2004b). Data Warehouse Efforts and Metadata Foundations. EDUCAUSE.

Thomas, Scott L. et al. (2005). "Weighting and Adjusting for Design Effects in Secondary Data Analyses." In *New Directions for Institutional Research*, no. 127. Survey Research: Emerging Issues. pp. 51-72.

Trosset, Carol and Weisler, Steven. (2010). "Using Longitudinal Assessment Data to Improve Retention and Student Experiences." In *New Directions for Institutional Research*, Assessment Supplement 2010. pp.79-88.

VCCS. (2010). "Perkins Core Performance Measures: Results and Targets 2008-98." Richmond, VA: Virginia Community College, Workforce Development Services.

Visher, Mary G. et al. (2012). "The Effects of Learning Communities for Students in Developmental Education: A Synthesis of Findings from Six Community Colleges." National Center for Postsecondary Research.

VSA. (2008). "Voluntary System of Accountability: Information on Learning Outcomes Measures." Learning Outcomes Technical Work Group.

Voorhees, Richard. A. and Milam, John. (2005). The Hidden College: Noncredit Education in the United States. Stephens City, VA: HigherEd.org.

Walpole, Mary Beth. (2008). "Emerging from the Pipeline: African American Students, Socioeconomic Status, and College Experiences and Outcomes." *Research in Higher Education*, 49. pp. 237-255.

WCET. (2012). "2010 Federal Regulations on State Approval of Out-of-State Providers." WCET Advance. Boulder, CO: WICHE Cooperative for Educational Technologies.

WIKI. (2012). "Open Source Databases Comparison." Encore Wiki document. Available online at: http://www.encorewiki.org/display/encore/Open+Source+Databases+Comparison.

Wikipedia. (2012a). "The Open Source Definition." Wiki document. Available online at: http://en.wikipedia.org/wiki/Oen_Source_Definition.

Wikipedia. (2012b). "Business Intelligence Tools." Wiki document. Available online at: http://en.wikipedia.org/wiki/Business_intelligence_tools.

Witt, Allison et al. (2012). "Dual Credit/Dual Enrollment and Data-Driven Policy Implementation: Reform Initiatives and Postsecondary Credential Attainment." Illinois Board of Higher Education, Illinois Education Research Council.

Wittstruck, John et al. (2002). "Linked Longitudinal Administrative Data Sets." Presentation at the 2002 NCES/IPEDS Network Conference in Washington, D.C. in May, 2002.

# Websites

| Website | Type | URL |
|---|---|---|
| Achieving the Dream | National initiative to improve community college student | |
| Common Education Data Standards (CEDS) | Federal program for data standards | http://ceds.ed.gov |
| Completion Arch | | http://completionarch.collegeboard.org |
| DELTA Cost Project Database | institution-level data for trends in costs | http://nces.ed.gov/ipeds/deltacostproject/ |
| | | |
| | Online data analysis application, NCES | http://nces.ed.gov/das/ |
| Education Resources Information Center (ERIC) | Federally-funded online resource for educational research | http://www.eric.ed.gov/ |
| Illinois Education Research Council (IERC) | Conducts research with SLDS for Illinois | http://www.siue.edu/ierc/ |
| Illinois P 20 Council | Makes recommendations about SLDS | http://www2.illinois.gov/gov/P20/Pages/About.aspx#what |
| IPEDS Data Center | Online data analysis applications, NCES | http://nces.ed.gov/ipeds/datacenter/ |
| National Crosswalk Service Center | Federal program | http://www.xwalkcenter.org |
| NCES CIP 2010 | Federal taxonomy of disciplines | http://nces.ed.gov/ipeds/cipcode/ |
| NCES IPEDS | NCES IPEDS site | http://nces.ed.gov/ipeds/ |
| NSF | NSF data collection site | http://www.nsf.gov/statistics/ |
| Postsecondary Electronic Standards Council | Federal program for data standards | http://www.pesc.org/ |
| PowerStats | Online data analysis application, NCES | http://nces.ed.gov/datalab/ |
| Privacy Technical Assistance | Federal program for privacy | http://ptac.ed.gov/ |

| Website | Type | URL |
|---|---|---|
| Center | | |
| SESTAT | Online data analysis application | http://www.nsf.gov/statistics/sestat/ |
| Source Forge | | http://www.sourceforge.net |
| State Authorization Reciprocity Agreement (SARA) draft | Draft of state agreement about serving non-resident online students | http://www.csg.org/NCIC/documents/WorkingDraft.pdf |
| State Data Systems and Privacy Concerns | Publication from Jobs for the Future | http://www.jff.org/sites/default/files/StateDataSystems.pdf |
| State Longitudinal Data Systems | Federal program | http://nces.ed.gov/programs/slds/ |
| The Consortium for Student Retention Data Exchange (CSRDE) | Higher education consortium | http://csrde.ou.edu/web/index.html |

## About the Author

John Milam earned a Ph.D. in higher education from the University of Virginia (1989) and has held institutional research and faculty positions at the University of Houston, West Virginia University, George Mason University, the University of Virginia, and Lord Fairfax Community College. He developed the GMU Data Warehouse in 1995; built websites for AIR, ASHE, and the ERIC Clearinghouse in the early years; and hosted *Internet Resources for Institutional Research* from 1995 to 2005. He's taught a number of workshops over the years about data applications on the web, including one on *Data Warehousing: An Introduction* at the 2012 AIR Forum.

CEO of HigherEd.org, Inc. until 2009, John led the development team for various online IPEDS data applications, including the Executive Peer Tool, Dataset Cutting Tool, Glossary, Data Analysis System, CIP 2010, and State Data Center; also conducting special IPEDS studies of derived variables, online software design, data allocation schemes, and noncredit data collection. Some of his other projects include the NPEC ANSWERS website, the Virginia Teacher Education and Licensure Data Warehouse, and the Common Data Set Exchange. He was the contractor for and co-authored the IPEDS Student Unit Record Feasibility Study in 2005 and has conducted several multi-state, longitudinal studies for NCHEMS and as part of a Lumina Foundation for Education grant through the University of Virginia about serving non-traditional students.

John has published about taxonomies, web applications for Institutional Research, and using national datasets. He is the co-author of a chapter in the Handbook for Institutional Research (2012) about Business Intelligence and Analytics and another about Building Cost Models.