1

This chapter opens with a literature scan examining the use of longitudinal pretest-posttest designs in college impact research. With that in mind, the authors present an example of how researchers can mis-estimate college impact if pretest-posttest designs are not used.

The Importance of Longitudinal Pretest-Posttest Designs in Estimating College Impact

Tricia A. Seifert, Ernest T. Pascarella, Sherri I. Erkel, Kathleen M. Goodman

In this chapter, we discuss the issue of research design in conducting inquiry on college impact and demonstrate the importance of longitudinal pretest-posttest designs in maximizing the internal validity of findings. We begin by discussing the strengths and weaknesses of different types of research design in the college impact literature (for example, cross-sectional, longitudinal with a pretest, and longitudinal without a pretest) and continue by presenting the results of a five-year review of research studies published in four major higher education journals: *Journal of College Student Development*, *Journal of Higher Education*, *Research in Higher Education*, and *Review of Higher Education*. Within this review, we present the percentage of college impact studies published in these journals from 2004 to 2008 using the aforementioned research designs. We then offer an empirical example based on data from the Wabash National Study of Liberal Arts Education (the Wabash Study). This example demonstrates the extent to which the effects of an independent variable can be speciously over- or underestimated when a pretest is not used in research. We conclude with a discussion of the implications for conducting internally valid studies of college impact.

Educational Research Design

Two basic quantitative research designs exist for measuring college impact: cross-sectional and longitudinal panel studies. Cross-sectional studies

Published online in Wiley Online Library (wileyonlinelibrary.com) • DOI: 10.1002/ir.368

collect data at one point in time from a group that can range in age, year in school, or type of institution attended. In estimating the impact of college experiences on student learning, cross-sectional designs attempt to statistically control for all confounding influences that may affect the relationship between the independent variable and outcome under examination. Intellectual ability from high school (for example, measured by high school grades or a score on an objective college entrance exam such as the ACT or SAT) is often used as the covariate from which a baseline of student academic ability is estimated. The one-time nature of the data collection eliminates the possibility of sample attrition and reduces costs associated with data collection, which are clear benefits. However, the challenge with cross-sectional design in college impact studies is that the covariates selected are usually convenient but imperfect and imprecise proxies for an actual baseline (Pascarella, 2006). The result from this imprecision is a less accurate estimate of the relationship under investigation.

Unlike cross-sectional design, longitudinal panel study design allows educational researchers the advantage of surveying the *same* sample of students at two or more moments in time. By following the same cohort of students, researchers can estimate how much an individual changed between time points. The principal benefit of panel study design is that it enables researchers to better understand participants' characteristics from the outset, allowing more accurate statistical control throughout the duration of the study (Gall, Gall, and Borg, 2006; Singleton and Straits, 2009).

Longitudinal panel study is further delineated by those studies that include a parallel pretest measure of the outcome under examination and those that simply collect data from the same cohort at two or more points in time. In college impact research focused on specific learning outcomes, panel designs with a pretest allow researchers to account for students' baseline in measuring the impact (or the value-added) of college experiences on their learning. Pascarella (2006) claims the data in longitudinal pretest-posttest studies furnish a better estimate of college impact than statistical manipulation from cross-sectional data. Astin's Input-Environment-Output model (1991, 2003) is also based on this notion that longitudinal study design offers stronger and more reliable data than cross-sectional design, particularly when the inputs include pretest measures of outcomes.

There are situations in which researchers may be interested in understanding the impact of college experiences on an outcome for which no parallel pretest exists; degree attainment is one such example. In this and similar instances, researchers have identified measures that serve as proxies for an actual pretest. For example, in measuring the influence of attending a selective institution on degree completion for minority students, Melguizo (2008) used theoretically based proxies such as students' precollege degree goal commitment and commitment to the institution to isolate the net effects of the relationship under investigation. Others have

used similar kinds of theoretically based proxies from earlier data collection points to better estimate the impact of college experiences and environments on the outcome of interest (see Arbono and Nora, 2007; Astin and Lee, 2003; Locks, Hurtado, Bowman, and Oseguera, 2008; Rhee, 2008; and Walpole, 2008).

Although panel studies are not without their problems—repeated measurement, participant dropout, and the possibility of sample bias are not trivial concerns—they enable researchers to collect and analyze data that yield more internally valid results than any other research design, with the exception of randomized experiments (Gall, Gall, and Borg, 2006; Pascarella, 2006). Pretest-posttest design ultimately permits researchers to concentrate on observing and identifying the changes that occur in each participant in a more reliable manner than any other design, allowing researchers to make stronger inferences from the sample to the target population (Yee and Neimeier, 1996). By taking into account students' precollege characteristics and pretest measures of desired educational outcomes, longitudinal pretest-posttest design better equips higher education researchers to estimate the genuine impact of college experiences on educational outcomes (Astin, 1991; Pascarella and Terenzini, 1991, 2005).

Review of College Impact Research. Some may critique our call for the use of longitudinal pretest-posttest designs to measure college impact outdated; researchers have long been employing more rigorous designs of college impact. It is true; Feldman and Newcomb (1969) issued such a call in their seminal work forty years ago. It is noteworthy that this call has been echoed fairly recently by Astin (2003) and Pascarella (2006). We tested the veracity of the "outdatedness" critique by reviewing five years of studies (2004–2008) published in four major higher education journals: *Journal of College Student Development*, *Journal of Higher Education*, *Research in Higher Education*, and *Review of Higher Education*. We found that college impact research, broadly defined, constituted 33 percent of the published articles. Of those college impact studies, 50 percent employed cross-sectional designs, typically using high school grades or scores on objective college entrance examinations as baseline proxies for student learning. Only a quarter of the college impact research published in these journals used a longitudinal design with a pretest. For those who assert that a call for more consistent use of pretest-posttest research designs in college impact research is passé and that the literature has already made these design adjustments, we offer the findings from this review as support to counter the claim. Longitudinal pretest-posttest designs may not be quite the exception they once were, but they certainly are not the rule. We present an empirical example to demonstrate the efficacy of longitudinal pretest-posttest designs in estimating college impact.

Empirical Example. The Wabash Study is one such project that has used a quantitative longitudinal pretest-posttest design within a broader

mixed-methods study to assess the impact of college on student learning. Within the quantitative portion of the Wabash Study, the research team collected data from students at three time points during their college career: at the beginning and end of their first year and again at the end of their fourth year. Designed to enrich educational stakeholders' understanding of the teaching practices, student experiences, and institutional conditions that promote liberal arts education, the Wabash Study does not assume the liberal arts are rooted in a particular set of arts and sciences disciplines or a certain kind of institution. Rather, the Wabash Study adopted an intentional "outcomes-based" approach to the liberal arts such that any institution could be considered a liberal arts institution so long as its aims included commitment to developing students' critical thinking, moral reasoning, leadership toward social justice, well-being, interest in and engagement with diversity, and interest in deep intellectual work (see http://www.liberalarts.wabash.edu/storage/Defining_Liberal_Education.pdf).

This perspective and the inclusion of various types of institutions in the Wabash Study shape our empirical example and the guiding research question: To what extent does including a pretest in the analytic model change the effects of institutional type on liberal arts outcomes, when statistically controlling for the student background characteristics often used as covariates in cross-sectional designs? If differences between institutional types on the posttest measures are reduced in magnitude or cease to be statistically significant when a pretest is included in the model, this indicates that the covariates often used in cross-sectional designs are insufficient proxies for a true baseline pretest. Any change in institutional type effects when a pretest is entered into the model supports our assertion that in order to accurately estimate the net impact of college experiences on student learning, both pretest and posttest measures on outcomes of interest are critical.

Sample. Our sample consisted of students from nineteen U.S. institutions, which ranged in geographic location, size, control, type, and patterns of student residence, participating in the 2006 cohort of the Wabash Study. We collected data at two times during fall 2006. Students who were interested in participating in the study completed either a paper or a web-based registration form providing information on demographic characteristics and high school involvement. Those who completed the registration form were sent a letter and email reminding them to attend an outcomes assessment session. At these sessions, we collected pretest data on students' orientation toward learning and career plans, and six outcomes associated with liberal arts education. Because of the cognitive demands of two of the outcomes (critical thinking and moral reasoning), the sample was divided, with each group completing only one of the two outcome measures. We also collected data from participants at the end of their first year of college. Students completed two college student experience surveys as well as posttest data on all outcome measures during the spring

2007 data collection. Students received $50 for participating in each data collection period.

Analyses. To answer our research question, we ran three models for each outcome. In the first model, we approached the analysis from a cross-sectional perspective for which student background characteristics are totally unaccounted. The first model included only the three dichotomous measures of institutional type (community colleges, regional institutions, and research universities with liberal arts colleges as the reference group) to predict posttest measures of educational outcomes. This is often done in large-scale comparison studies of graduation rate (see Astin and Oseguera, 2002). The second model assessed whether students' posttest measures of educational outcomes differed with the institutional type, controlling for student background characteristics. The controls for student background characteristics included gender, race or ethnicity (white vs. student of color), parents' educational attainment, family income, whether student has one or more dependents, high school racial composition, high school involvement, work for pay in high school, college entrance examination score (ACT, SAT, COMPASS), whether student had transfer credits, and whether student attended his or her first-choice college. This type of analysis models that typically used in cross-sectional designs in which student background characteristics are included as covariates and proxies for precollege student learning or ability. It is also used in longitudinal designs in which there is no precollege measure of the outcome. The third model assessed the influence of a pretest on estimates of institutional type, when student background characteristics have been controlled. This type of analysis models the longitudinal pretest-posttest design.

The dependent variables in these analyses were posttest measures representing liberal arts educational outcomes (see King, Kendall Brown, Lindsay, and VanHecke, 2007): the Openness to Diversity/Challenge Scale (Pascarella and others, 2005); the Miville-Guzman Universality Diversity Scale, short form (Fuertes and others, 2000); the Socially Responsible Leadership Scale, revised version (Appel-Silbaugh, 2005; Dugan and Komives, 2006); the Scales of Psychological Well-being (Ryff, 1989; Ryff and Keyes, 1995); the Need for Cognition (Cacioppo, Petty, and Kao, 1984); the Positive Attitude Toward Literacy Scale (Pascarella and others, 2005); the CAAP Critical Thinking Test (ACT, 1990); and the Defining Issues Test-2 (DIT-2; Rest, Narvaez, Thoma, and Bebeau, 1999). All analyses are weighted estimates such that the analytic sample was weighted to better approximate the population. All analyses account for students being clustered within schools and the effects that these nonindependent observations can have on the standard error of the estimates (Raudenbush and Bryk, 2002).

Results. The three models for each outcome displayed in Table 1.1 illustrate the difference in effects found under various analytic designs.

Table 1.1. Estimated Effects of Institutional Type on Liberal Arts Outcomes

	Community Colleges vs. LACs			*Regional Institutions vs. LACs*			*Research Universities vs. LACs*		
	I	*II*	*III*	*I*	*II*	*III*	*I*	*II*	*III*
Interest in and engagement with diversity:									
Openness to Diversity Scale	–.405***	–.132	–.04	–.182	–.195*	–.086	–.202	–.219**	–.10*
Miville-Guzman Universality Diversity Scale, diverse orientation total	–.453***	–.074	–.123*	–.179	–.213**	–.089*	–.136	–.164	–.041
Leadership toward social justice:									
Socially Responsible Leadership Scale total	–.148	.091	.016	–.034	.008	.005	.121	.064	.058
Well-being:									
Scales of Psychological Well-Being total	–.163	.061	.04	–.124	–.076	–.051	.080	.016	–.013
Interest in deep intellectual work:									
Need for Cognition	–.530**	.084	.073	–.350**	–.194*	–.106*	–.095	–.243**	–.067
Positive Attitude Toward Literacy Scale	–.582**	.002	–.055	–.369***	–.293***	–.144***	–.252	–.346***	–.130**
Critical thinking:									
CAAP Critical Thinking Test	–.868***	.044	–.059	–.60	–.188*	–.111*	.307	.081	.027
Moral reasoning:									
Defining Issues Test-2 (DIT-2), postconventional moral reasoning	–1.002***	–.268**	–.240***	–.518**	–.268	–.086	–.026	–.173	–.098*

Notes: The coefficients presented in the table can be interpreted as effect sizes. A one-unit change in institutional type (for example, liberal arts college to community college) results in a *y* standard deviation change. All coefficients are the unique effect of the institutional type (compared to liberal arts colleges) on the outcome controlling for gender, race and ethnicity, parents' educational attainment, family income, whether student has one or more dependents, high school racial composition, high school involvement, work for pay in high school, college entrance examination score. whether student had transfer credits, and whether student attended his or her first-choice college.
* $p < .05$; ** $p < .01$; ***$p < .001$.

Looking at the columns identified as model 1 in the table, if one compares the effect of attending various kinds of institutions on liberal arts outcome development simply on students' scores at the end of their first year in college, one may assert community colleges are largely ineffective compared to liberal arts colleges at promoting these outcomes in their students. Note that on six of the eight outcomes, community colleges have a negative effect on the outcome. However, if we look at the community college model 2 column, it becomes apparent those institutional type differences were almost entirely due to the differential background characteristics of students who attend community colleges compared to liberal arts colleges. With one exception for the community colleges, all effects from model 1 cease to be statistically significant in model 2. This is also the case for the postconventional moral reasoning score for regional institutions; a negative effect in model 1 ceases to be statistically significant in model 2.

Comparing model 2 to model 1 for the regional institutions and research universities, we find another interesting set of findings. With the exception of the difference between regional institutions and liberal arts colleges on the two measures of Interest in Deep Intellectual Work (that remain statistically significant across all models) and postconventional moral reasoning, we found no differences in the posttest measures of the outcomes between regional institutions, research universities, and liberal arts colleges in model 1. However, by including student background characteristics in model 2, we identified a number of statistically significant negative effects between regional institutions, research universities, and liberal arts colleges. In six of the sixteen analyses investigating differences between regional institutions and research universities (compared to liberal arts colleges), regional institutions and research universities were less effective in promoting development on these liberal arts outcomes, once we controlled for differences in student background characteristics. From this perspective, a strict cross-sectional analysis such as model 1 would fail to pick up critical differences in outcome development that we were able to detect by using a cross-sectional analysis that accounted for student background characteristics.

Finally, we examined the influence of including a parallel pretest to the analytic models. Model 3 demonstrates what is lost (and gained) in estimating college impact using a longitudinal pretest-posttest design. In nearly half of the model 3 analyses (eleven of twenty-four), inclusion of a pretest yielded findings different from those of the cross-sectional analyses in which we controlled for student background characteristics (model 2). Including the pretest to the analytic models did one of three things to the estimation of college impact: (1) it reduced the magnitude of the difference between institutional types typically by half, (2) it eliminated statistically significant differences between institutional types, or (3) it showed differences between institutional types not identified from model 2.

Of the eleven instances in which we found differences from model 2 to model 3, seven were findings in which the magnitude of the institutional type effect identified in model 2 was reduced dramatically. For example, the analysis of regional institutions compared to liberal arts colleges in promoting positive attitude toward literacy outcome suggested regional institutions were less effective than liberal arts colleges by nearly 0.3 of a standard deviation, controlling for student background characteristics. However, once a parallel pretest was entered in model 3, this effect was reduced to about 0.15 SD. In other words, once we took into account students' baseline on the pretest, the difference was reduced by half between regional institutions and liberal arts colleges in promoting a positive attitude toward literacy.

In two cases, inclusion of a parallel pretest in model 3 resulted in the difference between institutional types ceasing to be statistically significant. This occurred in the analysis comparing regional institutions to liberal arts colleges on the openness-to-diversity measure in which the nearly 0.2 SD difference disadvantaging regional institutions was reduced to nonsignificance in model 3. We found a similar result in the analysis between research universities and liberal arts colleges on the need for cognition outcome. The negative effect of attending a research university compared to a liberal arts college ceased to be statistically significant once we controlled for differences on the parallel pretest.

Finally, in two situations, model 3 identified institutional type differences not found in model 2 (the cross-sectional design controlling for student background characteristics). In both the postconventional moral reasoning analysis between research universities and liberal arts colleges and the MGUDS analysis between community colleges and liberal arts colleges, inclusion of a parallel pretest resulted in negative effects between these two institutional types and liberal arts colleges for the outcomes under examination. In other words, by controlling for student differences on the pretest, we were able to identify the institutional type differences on the posttest.

Discussion

The intention of this chapter is twofold: (1) to estimate the percentage of college impact studies, published in major higher education journals, that used longitudinal pretest-posttest designs; and (2) to supply an empirical example of the over- or underestimation that can result in college impact studies that do not use a longitudinal pretest-posttest design.

In our review, we found about 25 percent of the published college impact articles in four major higher education journals used a pretest-posttest design. Although beyond our scope, it seems highly likely that we would find at maximum 25 percent of the numerous assessment projects taking place across the country using a longitudinal pretest-posttest design

to estimate the impact of any number of educational programs or services on student learning. To assert that longitudinal pretest-posttest designs are widespread in the literature and in practice is unsupported by our review. In light of the findings from our empirical example, when one couples the sizable percentage of published research studies with the vast number of smaller-scale institutional assessment efforts using cross-sectional designs to measure college impact on specific learning outcomes, the likely inaccurate attributions of curricular and co-curricular programs to institutional quality or effectiveness (whether these are positive or negative) become apparent.

The fact that we found such variation in the findings based on inclusion of a pretest is noteworthy. 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. Although it is quite common for cross-sectional (and some longitudinal) studies of college impact to collect and statistically control for student background characteristics (for example, gender, race and ethnicity, SES, ACT or SAT score), our analyses showed that even with such background influences controlled, including a pretest in the model resulted not only 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. Although our current example focuses on differences based on institutional type (attending a liberal arts college versus a community college, regional institution, or research university), similar results would be conceivable in estimating the net impact of any experience or intervention into which students self-select.

Our results show a clear example of selection bias. In other words, according to the type of institution students attended, they differed in statistically significant ways on the baseline measure—the pretest. The ultimate result of such significant selection bias on the pretest, for which cross-sectional designs cannot directly adjust, is increased likelihood of a confounded estimate of the impact of the educational program or service being investigated on the outcome. This confounded estimate would likely result in a researcher erroneously reporting a differential impact (or value-added) of a specific experience on student learning. In our example, we focused on institutional type, but these same results would occur if the independent variable were participation in a service-learning course or study abroad. The bottom line is that, absent the possibility of randomized experiments, longitudinal pretest-posttest designs are the best way to estimate what students are learning and how they are changing.

Given external stakeholders' growing focus on institutional accountability of student learning (Miller, 2007; Hersh, 2007; U.S. Department of Education, 2006) and internal stakeholders' demand for institutional

improvement and sound resource stewardship (Upcraft and Schuh, 2001), institutional researchers have an opportunity to aid their institutions, and postsecondary education in general, by conducting assessment in a manner that best facilitates data-driven decision making. Along with others (Astin, 2003; Astin and Lee, 2003; Pascarella, 2006), we recognize that longitudinal panel studies are not only time consuming but costly and difficult to conduct. However, in an accountability era known for demonstrating the value that programs and services add to student learning, we assert that longitudinal pretest-posttest panel designs yield the most internally valid results and the most accurate estimate of college impact. There is no substitute for the "gold standard" that longitudinal pretest-posttest studies furnish in accurately assessing how students learn and change.

We clearly recognize that longitudinal, pretest-posttest designs put a greater burden on institutional researchers in terms of time, effort, and resources than do cross-sectional studies. However, as our analyses have shown, with cross-sectional studies one almost always runs the substantially greater risk of mistaking a program effect for what is really a recruitment or self-selection effect. Because of this, we caution institutional researchers who use cross-sectional designs to purport program effectiveness to reiterate the inherent limitations of these designs in making evidence-based recommendations to internal and external stakeholders. We suggest institutional researchers use the findings from this empirical example to build the case for allocating institutional resources such that more internally valid studies examining the impact of curricular and co-curricular programs and services can be conducted. Ultimately, one must be willing to make a choice between longitudinal pretest-posttest studies, which generate findings with greater internal validity but are more difficult to conduct, and easier-to-conduct cross-sectional studies that increase the likelihood of confounded findings.

References

American College Testing Program (ACT). "Report on the Technical Characteristics of CAAP; Pilot Year 1: 1988–89." Iowa City: ACT, 1990.

Appel-Silbaugh, C. "*SRLS Rev: The Revision of SRLS*." College Park, Md.: National Clearinghouse for Leadership Programs, 2005.

Arbona, C., and Nora, A. "The Influence of Academic and Environmental Factors on Hispanic College Degree Attainment." *Review of Higher Education,* 2007, *30*(3), 247–269.

Astin, A. W. *Assessment for Excellence*. San Francisco: Jossey-Bass, 1991.

Astin, A. W. "Studying How College Affects Students." *About Campus,* 2003, *8*(3), 21–29.

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

Astin, A. W., and Oseguera, L. *Degree Attainment Rates at American Colleges and Universities: The Latest*. Los Angeles: Higher Education Research Institute, 2002.

Cacioppo, J. T., Petty, R. E., and Kao, C. F. "The Efficient Assessment of Need for Cognition." *Journal of Personality Assessment,* 1984, *48*(3), 306–307.

Dugan, J. P., Komives, S. R., and Associates. *Multi-Institutional Study of Leadership: A Guidebook for Participating Campuses*. College Park, Md.: National Clearinghouse for Leadership Programs, 2006.

Feldman, K. A., and Newcomb, T. M. *The Impact of College on Students*. San Francisco: Jossey-Bass, 1969.

Fuertes, J. and others. "Factor Structure and Short Form of the Miville-Guzman Universality-Diversity Scale." *Measurement and Evaluation in Counseling and Development*, 2000, *33*, 157–169.

Gall, M. D., Gall, J. P., and Borg, W. R. *Educational Research: An Introduction (8th ed.)*. Boston: Allyn and Bacon, 2006.

Hersh, R. "Going Naked." *Peer Review,* 2007, *9*(2), 4–8.

King, P. M., Kendall Brown, M., Lindsay, N. K., and VanHecke, J. R. "Liberal Arts Student Learning Outcomes: An Integrated Approach." *About Campus,* 2007, *12*(4), 2–9.

Locks, A. M., Hurtado, S., Bowman, N. A., and Oseguera, L. "Extending Notions of Campus Climate and Diversity to Students' Transition to College." *Review of Higher Education,* 2008, *31*(3), 257–285.

Melguizo, T. "Quality Matters: Assessing the Impact of Attending More Selective Institutions on College Completion Rates of Minorities." *Research in Higher Education,* 2008, *49*, 237–255.

Miller, R. *Assessment in Cycles of Improvement: Faculty Designs for Essential Learning Outcomes*. Washington, D.C.: Association of American Colleges and Universities, 2007.

Pascarella, E. T. "How College Affects Students: Ten Directions for Future Research." *Journal of College Student Development,* 2006, *47*(5), 506–520.

Pascarella, E. T., and Terenzini, P. T. *How College Affects Students*. San Francisco: Jossey-Bass, 1991.

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

Pascarella, E. T. and others. "Liberal Arts Colleges and Liberal Arts Education: New Evidence on Impacts." *ASHE Higher Education Report*, *31*(3). San Francisco: Jossey-Bass, 2005.

Raudenbush, S. W., and Bryk, A. S. *Hierarchical Linear Models: Applications and Data Analysis Methods (2nd ed.)*. Thousand Oaks, Calif.: Sage, 2002.

Rest, J. R., Narvaez, D., Thoma, S. J., and Bebeau, M. J. "DIT2: Devising and Testing a Revised Instrument of Moral Judgment." *Journal of Educational Psychology,* 1999, *91*, 644–659.

Rhee, B. S. "Institutional Climate and Student Departure. A Multinomial Multilevel Modeling Approach." *Review of Higher Education,* 2008, *31*(2), 161–183.

Ryff, C. D. "Happiness Is Everything, Or Is It? Explorations on the Meaning of Psychological Well-being." *Journal of Personality and Social Psychology,* 1989, *57*(6), 1069–1081.

Ryff, C. D., and Keyes, C.L.M. "The Structure of Psychological Well-Being Revisited." *Journal of Personality and Social Psychology,* 1995, *69*(4), 719–727.

Singleton, R. A., and Straits, B. C. *Approaches to Social Research (5th ed.)*. New York: Oxford University Press, 2009.

Upcraft, L., and Schuh, J. *Assessment Practice in Student Affairs: An Applications Manual*. San Francisco: Jossey-Bass, 2001.

U.S. Department of Education. "A Test of Leadership: Charting the Future of U.S. Higher Education." 2006. Retrieved Apr. 20, 2007, from http://www.ed.gov/about/bdscomm/list/hiedfuture/reports/final-report.pdf.

Walpole, M. "Emerging from the Pipeline: African American Students, Socioeconomic Status, and College Experiences and Outcomes." *Research in Higher Education,* 2008, *49*, 237–255.

Yee, J., and Niemeier, D. "Advantages and Disadvantages: Longitudinal vs. Repeated Cross-section Surveys." *Project Battelle 94(16). FHWA, HPM-40*, 1996. Retrieved Apr. 25, 2007, from http://ntl.bts.gov/lib/6000/6900/6910/bat.pdf.

TRICIA A. SEIFERT is an assistant professor in the Department of Theory and Policy Studies at the Ontario Institute for Studies in Education at the University of Toronto.

ERNEST T. PASCARELLA is the Mary Louise Petersen Professor of Higher Education and co-director in the Center for Research on Undergraduate Education at the University of Iowa.

SHERRI I. ERKEL is a doctoral candidate in the Student Affairs Administration and Research program at the University of Iowa.

KATHLEEN M. GOODMAN is a visiting instructor in the Student Affairs and Higher Education program at Miami University.