Dear NEAIR Friends and Colleagues,

The 2001 annual conference of the Northeast Association for Institutional Research was held from November 17th to the 20th at the beautiful Hyatt Regency Hotel in Cambridge, Massachusetts. The inspiration and spirit of the conference is captured in the theme, "Institutional Research: Leadership through Excellence." Some 246 conference attendees shared in the celebration of our association and profession. With many innovative features, the conference explored how institutional researchers can become leaders in higher education by effecting policy changes through quality research studies.

Highlights of the conference program include a stimulating opening keynote address, *Increasing Institutional Researchers' Influence* by Doctor Allan Cohen, Distinguished Professor of Global Leadership at Babson College. Jamie Merisotis, President of the Institute for Higher Education Policy, offered a thought-provoking plenary session, *Turning Higher Education Research into Results: A Policy Perspective*. The program also offered two excellent panel presentations - *Beyond Traditional IR: Allies in Higher Education Research*, convened by John Pryor, and *Careers in IR: (and beyond?)*, moderated by Jennifer Brown. These events, and an impressive array of poster sessions, research papers, table topics, workshares, and workshops contributed to a very enriching program. The Proceedings represent some of the intellectual content and insights shared during the conference.

The 28th annual conference marked the culmination of many months of creative planning and dedicated service by our Boston conference planning team. Our Local Arrangements Chair, Bea Frain, and Program Chair, Kelli Armstrong, deserve special recognition and admiration for their tireless team effort. Bea Frain's elegant taste, extraordinary organization, and attention to detail were evident in every aspect of local arrangements, especially in the delicious dinners and receptions, beautifully arranged conference settings and seamless conference scheduling. Thank you, Bea, for a superb job!

As reflected in these proceedings and throughout the conference, Kelli Armstrong successfully created a high caliber program with rich offerings appealing to new members of NEAIR as well as to seasoned professionals. Kelli's qualities - her sensitivity to each person, flexibility with competing demands, and creative vision for the profession - contributed to a unique and memorable program.

In addition to recognizing the conference chairs, I would also like to thank Ron Bentley who provided the wonderful jazz entertainment; Kathy Keenan who contributed her considerable artistic talents to the conference program; each member of the steering committee who dedicated hours to planning the conference; Marge Wiseman, Phyllis Fitzpatrick and our many volunteers who offered time and service during the conference; and each person who attended the conference. On behalf of all NEAIR members, I would like to express gratitude to Beth Simpson, our Membership Secretary, for her service during the year and for invaluable support with registration during the conference.

Finally, I want to express my gratitude to Marianthi Zikopoulos, our Publications Chair, for the many hours she has spent in preparing this exceptional document. Current and future members of NEAIR will benefit from her careful work and professional expertise. In content and form, these Proceedings reflect the goal of our conference to influence policy through excellence in research.

Anne Marie Delaney
NEAIR President, 2000 - 2001

2000 - 2001 Steering Committee

Officers

President
Anne Marie Delaney
Babson College

President Elect
Jim Trainer
Villanova University

Past President
Fred Volkwein
Pennsylvania State University

Secretary
Eleanor Swanson
Monmouth University

Treasurer
Mary Ann Coughlin
Springfield College

Members at Large

Jim Fergerson
Bates College

Martha Gray
Ithaca College

Heather Kelly Isaacs
University of Delaware

John Pryor
Dartmouth College

Steve Thorpe
Drexel University

Robert K. Toutkoushian
University System of New Hampshire

2001 Conference Chairs

Program Chair
Kelli Armstrong
University of Massachusetts-Boston

Local Arrangements Chair
Bea Frain
Massachusetts Institute of Technology

Publications Chairs
Marianthi Zikopoulos
New School University

Membership Secretary
Beth Simpson
HEDS Consortium

TABLE OF CONTENTS

*****This paper received the 2001 Best Paper Award**

THE ROCKY ROAD TO GRADUATION:
AN ACADEMIC CAREER FLOW MODEL
FOR TRACKING STUDENT PROGRESS AT COMMUNITY COLLEGES

Karl Boughan
Director of Institutional Research
Prince George's Community College

INTRODUCTION

During the last several decades, educational researchers have tended to answer the question *"Why do some college students succeed and others fail?"* by creating causal/predictive models of academic outcomes resting mainly on social psychological interpretations of collegiate behavior. Seminal examples are Tinto's social/academic integration model (1987, 1993), Astin's social input/academic output model (1977, 1993) and Bean's economy of organizational incentives model (1982). Unfortunately, while this approach may be appropriate in the development of a general theory of student academic progress, it provides educational planners with very little practical information.

When college administrators ask the crucial question, they are likely to mean *"Where in the academic process are we losing students?"* (e.g., during basic skills remediation, at the "gatekeeper" entry course point, while students are trying to satisfy their core curriculum requirements) and *"Why are they dropping out at these points?"* (e.g., poor academic performance, attendance fatigue, financial and family pressures). These practical queries are best answered by means of heuristic/descriptive models of student flow configured in terms of concrete academic processes. This paper presents such a model — one which maps enrollment flow in detail through (and out of) the academic process of a large, suburban community college, according to which student academic career phases coordinate with instructional program sequences. Its utility as a tool in enrollment management and instructional planning is illustrated in a model-guided six- year outcome analysis of fall entry cohort data.

Student Flow Models in Institutional Research

At the core of each student flow model is a map explicitly portraying a network of academic process streams through which students move over time and a coordinate set of flow measures which gauge movement volume at and between stream junctures. Promoted by the National Center for Higher Education Management Systems, which wedded projective techniques to this basically descriptive methodology, flow models enjoyed modest popularity among institutional researchers in the 1970s and 1980s (e.g., Johnson, 1974; Young, 1982). NCHEM-inspired models were created mainly as practical planning instruments for predicting

such things as future course demand, degree major enrollment, and facilities usage. Beyond primitive term-to-term retention analysis, however, little attention was paid to the potential of flow models as tools of student outcomes research. This oversight has been somewhat amended by the current spate of published research blending flow modeling with event probability modeling (e.g., Ronco, 1995; Kelly, 1996; Guerin, 1997; DesJardins and Moye, 2000) to explore the timing of types of enrollment outflows, and also by a number of recent flow-based studies centering on the outcome effects of particular academic process components (see St. John et al., 1991; Andrade 1999; Kramer and LaMar, 2000; Schartman and Rhee, 2000).

From a comprehensive planning perspective, the main problem with the former studies is that they focus too narrowly on outcomes timing issues, reducing student academic careers to a simple traversal of semesters. Academic process structure and operation, the factors most interesting to college planners, are either ignored in model construction or appear only as auxiliary control factors. The process component studies, on the other hand, tend to be more descriptive of process structure and operation and are therefore more useful to assessment and planning efforts, but their limited scope provides only small glimpses rather than the grand vista needed for full-scale *strategic* planning. The investigative lens for capturing the panorama properly must give a field of vision both wide and deep for obtaining an encompassing yet textured image of the whole academic process.

Constructing a Student Academic Career Flow Model

A student flow model capable of locating and measuring major enrollment management problems must be crafted to reflect the fundamental linked phase structure of the instructional process as it affects and is experienced by those undergoing it. In the case of most community colleges, these are easy to identify.

Commonly, community colleges, as non-selective admissions institutions, have needed to craft comprehensive "developmental" pregrams to bring large proportions of their student bodies deficient in basic English language and mathematical skills up to acceptable levels of credit course-taking preparedness. Then, for completing remedial students and those entering as college-ready, a "general education" phase begins. This usually consists of two subphases: First comes a set of "gateway courses" in English composition and postsecondary mathematics (designed to hone the key communications and analytic skills students need for success in tackling subsequent college subjects). This is followed by general education proper — a regimen of entry-level courses in a prescribed variety of fields (representing the range of human knowledge and intellectual endeavor and providing the grounding for the in-depth study of particular disciplines). Given the educational logic distinguishing them, in this study we will treat gateway courses and the general education residuum as two separate, closely linked academic phases. Lastly, in cumulative phased learning terms, students round off their academic careers by developing a level of expertise in one academic discipline or occupational area. This is accomplished by working through a particular degree program of specialized courses — the

"degree concentration" phase.

In addition to the above four instructionally defined phases (developmental, gateway, general education and degree concentration), one can posit a chronologically defined phase cutting across the others. It is no news that the great majority of U.S. postsecondary educational institutions experience disproportionately high early enrollment attrition rates. Prince George's Community College is no exception in this regard, typically losing around a quarter of its first-time college entry students between the initial fall and spring semesters. Many studies interpret this almost universal phenomenon as "college shock" — new student difficulties of adjustment to the unexpected rigors of college study and the alianness of college folkways. Since such a social psychological syndrome would constitute independent negative enrollment force acting independently of any academic process enrollment effects, we felt it prudent somehow to represent it in our student flow model as a separate dimension. In practice, this required splitting the developmental course-taking phase into preliminary first fall retention and remainder term subphases, and inserting a full non-developmental student first-fall retention phase before the gateway course-taking phase.

The fact, however, that the instructional apparatus found at most community colleges, including PGCC, was meant to work as a progression of course types (from general skills to general knowledge to specialized knowledge instruction) does not, of course, guarantee that it actually functions this way. Indeed, the PGCC catalog presents the developmental, gateway, general education and degree concentration study components as simple graduation requirements to be completed in no stipulated order. Also, practical considerations of course load and class scheduling often necessitate cross-phase course enrollments.

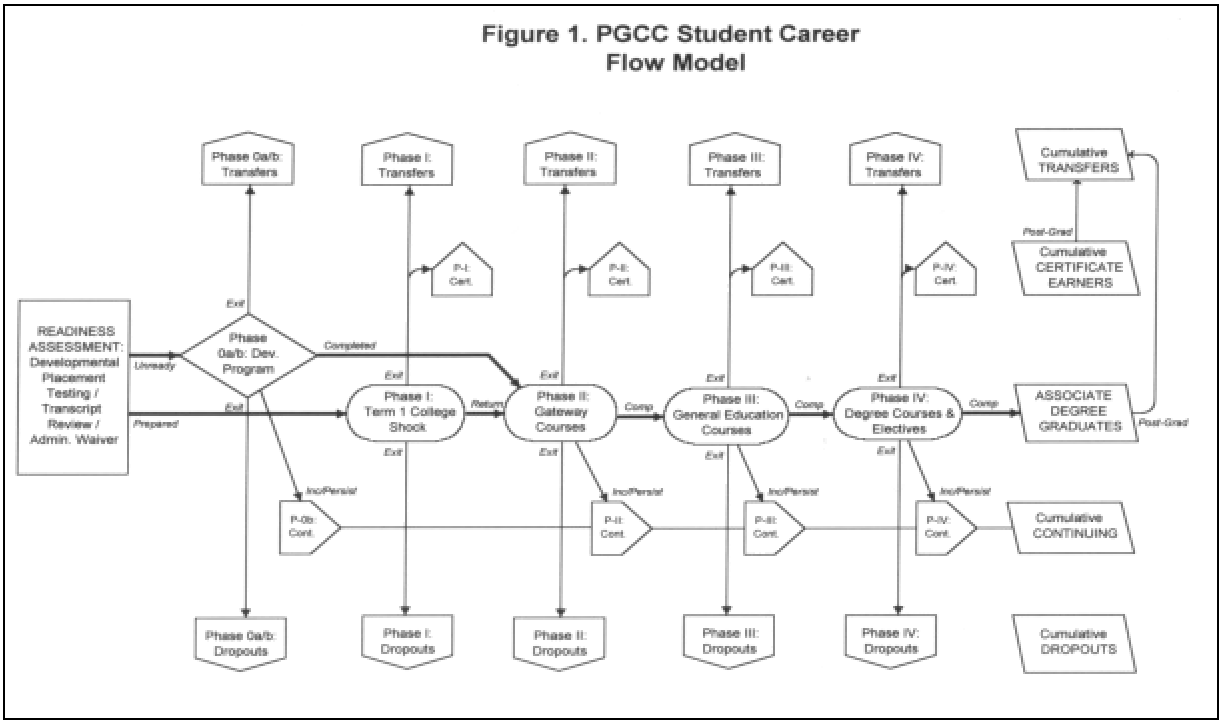
Even so, we find that most student course-taking does tend overall to follow the intended trajectory. The reasons are several: (1) Most students avoid random course-taking and seem naturally to order their academic careers in rough conformity with the education progression implicit in the college catalog, although they probably think more in terms of an "easy courses first" strategy than learning theory (the challenge posed by freshman English always seems to take them by surprise); (2) Academic counselors and faculty constantly prompt students to behave in reasonable compliance with the preferred study plan; (3) At PGCC, a fairly rigorous system of credit course pre-requisites promotes sequential phase course enrollment (for example, all gateway and most general education courses require prior completion of relevant remedial course work); (4) Students who habitually attempt advanced courses without proper grounding will naturally make little progress toward graduation and run a high risk of early dropout.

All of the above act to shape a genuine system of sequenced instructional phases in which meaningful course work within an instruction phase is essentially restricted to students having completed all prior phases course work. The truth of this is strongly suggested by an attempted credit hour analysis we carried out on the Cohort 1995 data: For the first phase transition, students requiring but not finishing their remedial programs attempted only 2.2 credit hours worth of gateway courses, on average, compared with 8.3 gateway credit hours for non-

developmental and developmental completing students. For the second transition, the mean attempted number of general education credit hours attempted by gateway phase non-completers came to only 5.1 against 13.9 attempted hours for completers. And in the third phase transistion case, general education non-completers averaged only 7.2 degree concentration credit attempts while general education completers accumulated an average of 36.1 attempted hours. Furthermore, only 66 out of 2,239 1995 Cohort students ever complete an instructional phase out of turn.

These findings lend empirical credence to our basic approach. Enrollment flow, by and large, does seem to follow a sequence of instructional phases, at least at PGCC, and therefore may be modeled accordingly. Unfortunately, the research also made clear that such phase-based flow was imperfect. Although the volume seemed too small to be of real consequence in the academic careers of most Cohort 1995 students, out-of-phase sequence course work did occur often enough to present us with a vexing technical problem. In a properly constituted flow model, students must work *completely* through one flow phase *before* proceeding to the next. Our solution to this problem was to employ the following model-building rule: Formally discount any out-of-phase course work and treat *the first uncompleted instructional phase* as solely responsible for all sub-graduation outcomes. This modeling rule forces only the elimination from the model of trivial course-taking, and reinforces the prime purpose of the flow model, which is to identify the *root sources* of enrollment outflow. The resulting academic career model of student is displayed in Figure 1.

Model flow, as shown, progresses from left to right, with the central set of linked ovals representing the main progress-toward-degree stream; all other shapes shown represent non-degree track circumstances and outcomes, cumulative six year outcomes or associate degree attainment. Flow begins with a college readiness assessment shunt (rectangle) which divides all newly admitted students into two streams according to the results of developmental placement testing and supplementary methods of basic college skill evaluation. Students found requiring remediation here are entered into one or more pre-credit programs for reading, English and math skills development (diamond-indicated Phase 0). Internally, the developmental phase is divided into an initial term sub-phase (0a) for gauging the enrollment effects of "college shock" and a post-first fall Sub-phase 0b (not separately shown). In the meantime, those requiring no remediation proceed directly to the their first term of credit course work (Phase 1, the first oval), and surviving that, return the next spring to continue their gateway course work (Phase 2) in entry-level English composition and pre-calculus mathematics (10 credit hours). Joining them for Phase 2 study are the Phase 0 students who completed their developmental programs and are now vetted as credit-course-ready. Thereafter, phase-completing students go on to the general education Phase 3 (normally 16 credit hours in five subject areas ? humanities, social science, physical/biological science, computer literacy, and cultural diversity), and if successful here to Phase 4 degree program course-taking (35-38 credit hours worth of degree department and elective courses) and possibly graduation with an associate degree (end-of-process parallelogram).



The downward flows branching off from the main stream depict the "negative" or suspended outcomes associated with each instructional phase. The falling perpendicular arrows represent phase non-completers exiting PGCC without any apparent academic attainment (dropouts) and the down-slanting arrows portray non-completing non-attainers still in attendance after six years (continuers). Upward branching flows off the main phase stream indicate two varieties of "positive" attrition — non-completing students leaving the college after receiving occupational certificates (a kind of early, minor graduation) and those transferring without a degree from PGCC.

As already mentioned, the model was operationalized by means of a six year transcript analysis of the academic careers of the cohort of Fall 1995 new native entrants (N=2,239). The initial cohort database, derived from official student records, contained exhaustive term-by-term information on developmental testing, placement and program progress, credit course enrollment and performance, degree program and achievement, and attendance and withdrawal patterns. In addition, this material was supplemented by full-coverage transfer behavior data supplied by the National Student Clearinghouse, assuring an extremely high rate of accuracy in the separation of transfer from dropout outflows.

Student developmental phase history was captured by means of standard developmental status transcript flags, thoroughly cross-checked with original placement test scores, developmental course performance, and the presence of credit course prerequisite waivers which might retroactively affect official remedial status. Gateway and general education course-taking were modeled in accordance with 1995 college catalog specifications. However, participation in the vastly varying degree concentration

phase was not independently modeled. Here we simply applied the survivor principle: Any student completing all prior phases was taken automatically to be a participant in the final phase. This worked out well in practice, since no prior phase survivor proved after six years to have accumulated less that 27 earned credit hours (26 earned credit hours are required for completion of phase 3).

Model Assessment and Findings

In the remainder of this paper, we will explore how well the PGCC academic career flow model performed when put into operation with actual student transcript data, and what its behavior had to teach our program planners and enrollment managers. A well-functioning student flow model designed as a general-purpose device for describing enrollment through-put and diagnosing enrollment problems should be able to fulfill multiple needs. At a minimum, it would be capable of delineating the basic unfolding pattern of student enrollment histories, of pinpointing and assessing the seriousness of enrollment problems by academic career phases, of suggesting reasons for the enrollment erosion discovered, and of helping college administrators strategize programmed responses.

Figure 2.Cohort 1995 Six Year Cumulative Student Flow by Academic Career Phases

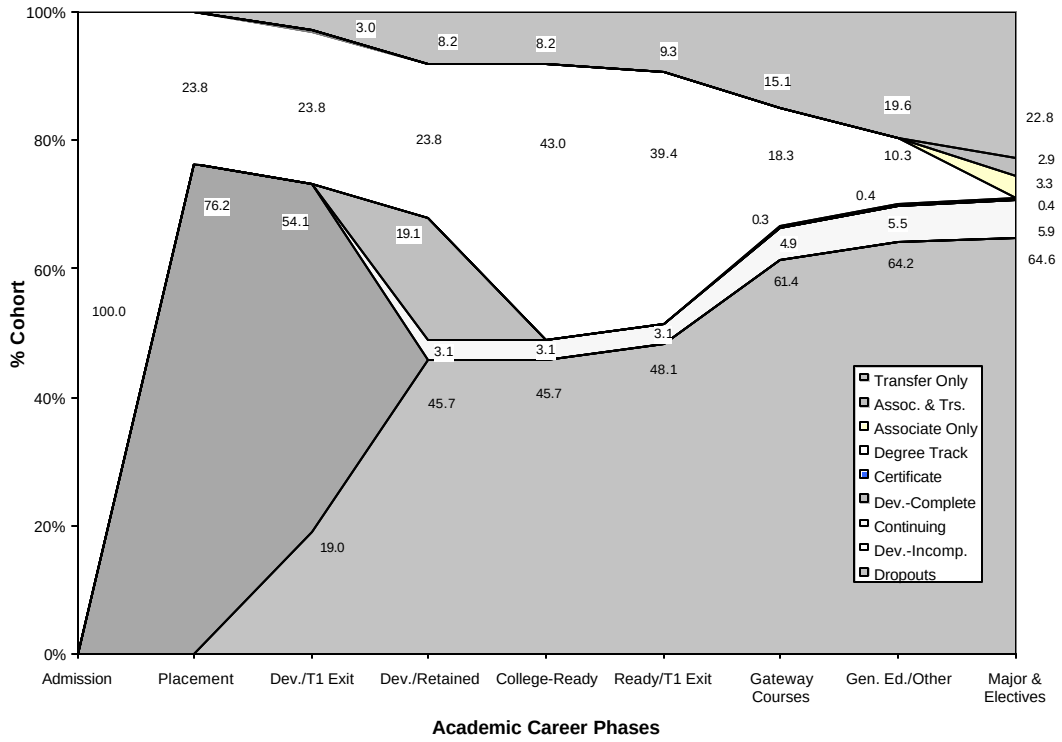


Figure 2, a stacked area chart, shows the model at work in portraying overall Cohort 1995 six-year outcomes flow. The area shadings identify the main retention-to-degree stream (white) and various categories of enrollment outflow. Area proportions (percent of total cohort) represent cumulative main

stream or outflow volumes at each successive academic career phase point. In the admissions pre-phase, 100 percent of Cohort 1995 students are shown as formally on degree track, changing quickly to 24 percent as the skills assessment pre-phase almost immediately shunts the bulk of new students (76 percent) into the remediation pre-credit pool. At the Phase 0a point (first term developmental course work), dropout outflow begin with a gush of 19 percent, surging to a full 46 percent in Phase 0b (remaining term developmental course work). The dispersal of the remediation pool becomes total with the formation of a developmental completer pool (19 percent) and the start-up of the continuing student sub-stream and transfer student outflow (both 3 percent). Then, at the credit course-ready collection point, the developmental completion pool empties back into the main stream, swelling it to a volume of 43 percent. From here forward, the pattern reduces to a simple matter of progressive main stream drainage up till graduation when the cohort diminishes to near zero (6 percent continuing students). Dropout outflow increases hardly at all at the non-developmental student first term point (Phase 1) , but spurts to 61 percent during gateway course-taking (Phase 2), thereafter climbing slowly to a 65 percent finale (degree concentration Phase 4). More importantly, the latter interval between Phases 2 and 4 sees a healthy growth in transfer-only outflow, reaching 23 percent in the end.

Table 1. Fall 1995 Entry Cohort Six Year Outcomes (N=2,239)

	Whole Cohort	***Outcomes Subsample***
Positive Outcomes	**29.6**	**100.0**
*All Transfers***	*25.7*	*87.0*
*All Transfers to 4 Year Schools***	*19.1*	*64.4*
*All Academic Awards***	*6.7*	*22.8*
Transfer to 2 Year School	6.7	22.6
Transfer to 4 Year School	16.1	54.4
Assoc. Degree + 4 Yr Transfer	2.9	10.0
Associate Degree Only	3.3	11.3
Occupational Certificate Only	0.4	1.5
Continuing/No Award or Transfer	**5.9**	**100.0**
Dropouts (by Phase)	**64.6**	**100.0**
*All Post-1st Term College Ready***	*17.3*	*26.8*
*All Developmental Dropouts***	*44.8*	*69.4*
*All 1st Term Dropouts***	*21.5*	*33.3*
P4-Degree Concentration Courses	1.3	1.9
P3-General Education Courses	2.8	4.3
P2-Gateway Courses	13.3	20.5
P1-1st Term Only Non-Developmental	2.5	3.8
P0b-Post –1st Term Developmental	25.8	40.0
P0a-1st Term Only Developmental	19.0	29.5

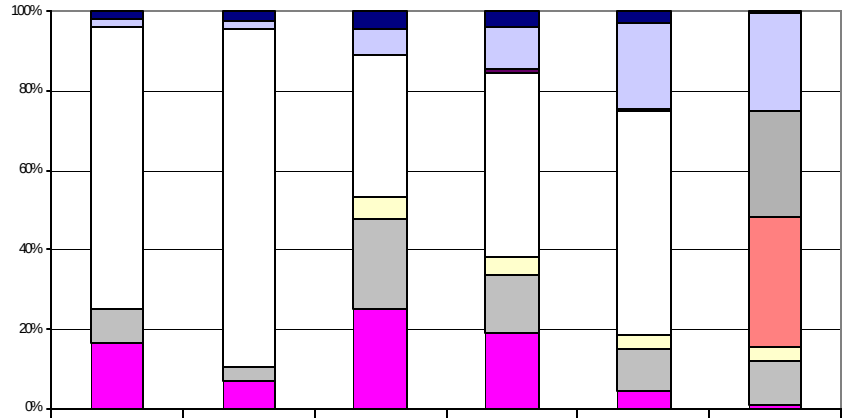
*60 or more cumulative earned credit hours after six years
**Overlapping summary outcome categories; *All Academic Awards* includes certificates; *All Post-1st Term College Ready Dropouts* includes Developmental Completers

Table 1 restates the results of six years of Cohort 1995 study in a final outcomes format, utilizing the student flow pattern findings to expand the dropout category (exiting PGCC without either award or transfer) so that it reflects occurrence by study phase. This application of the flow data turns an

otherwise ordinary outcomes report into a clear and precise means for locating enrollment erosion in academic process space and for assessing the relative seriousness of the problem represented by each outflow point.

According to the table, 30 percent of the students attained a "positive result" (transfer or some sort of academic award) after six possible years of PGCC attendance, while the bulk transfer without a degree (23 percent). Most of these were proper transfers to four-year schools (16 percent), but a significant minority made a lateral move to another community college (7 percent). Only 7 percent actually persisted all the way to graduation (almost all of these were associate degree earners). Rounding out non-dropout results, 6 percent turned out to students still attending (Spring 2001) and working on at least one outstanding graduation requirement. This left almost two-thirds (65 percent) in the dropout classification. By far, the weightiest dropout subset were students who got stuck at the developmental phase (45 percent of the whole cohort; 69 percent of all dropouts). The next biggest contributor to the dropout pool proved to be the gateway phase ; P2 exiting non-achievers represented 13 percent of all students and 21 percent of all dropouts. Only 4 percent of cohort students left defeated from either of the two advanced career phases (6 percent of all dropouts). Finally, we may note that, of the two first-term attrition phases, the developmental subphase produced a far greater proportion of dropouts (19 percent of the cohort; 30 percent of all dropouts) than the non-developmental student phase (only 3 percent of all students and 4 percent of all dropouts).

Figure 3. Individual Phase Outcomes (Stacked Bar % of Phase Participants)



	% P-0a: Dev. T1 Shock	% P-1: Non-Dev. T1 Shock	% P-0b: Dev. Retained	% P-2: Gateway Courses	% P-3: Gen Ed Remainder	% P-4: Degree Concentration
2Y TRANSFERS ONLY	1.9	2.6	4.7	3.9	2.9	0.4
4Y TRANSFERS ONLY	2.1	1.9	6.5	10.9	21.5	24.7
ASSOC+TRANSF (4Y)	0.0	0.0	0.0	0.0	0.0	26.8
ASSOCIATES ONLY	0.0	0.0	0.0	0.0	0.0	32.5
CERTIFICATE ONLY	0.0	0.0	0.0	0.9	0.5	0.0
PHASE COMPLETERS	71.1	85.2	35.3	46.3	56.5	0.0
CONTINUERS	0.0	0.0	5.8	4.4	3.4	3.5
DROPOUTS-G00D PERF.	8.7	3.6	22.5	14.5	11.0	11.3
DROPOUTS-P00R PERF.	16.3	6.7	25.2	19.1	4.2	0.9

A third way of looking at our flow data is to examine outflow types and rates phase by phase. Since each phase represents exposure to a different set of courses and subject matter, presenting different types and levels of academic challenge, individual phase outcomes analysis contextualizes student enrollment behavior as responses to process-driven variations in the learning experience. This in turn may lead to the formulation of more fitting intervention strategies and more effective instructional reforms. Figure 3 embodies our phase outcomes analysis of Cohort 1995 behavior in stacked bar chart form. Each bar reflects the enrollment flow consequences of participating in the course work required in a particular phase, and therefore is based solely on data for students actually reaching that phase in their academic careers.

Phase sample rate for all phase-relevant flow categories are shown (the full range includes degree-less transfer, transfer with an associate degree, associate degree only attainment, occupational certificate only attainment, phase completion and advancement to the next phase, continued attendance without phase completion, and dropping out). In addition, dropout and transfer flows are shown divided into sub-flows. Two sorts of dropouts are distinguished ? those dropping out with poor academic records and exiters whose level of academic performance seemed sufficient to assure eventual graduation. Classification was based on placement on an additive scale involving three performance variables (Cumulative GPA, proportion of courses passed, and proportion of semesters in official good academic standing). The transfer division was based on whether the move involved a four-year school (progressive transfer) or a shift to another community college (lateral transfer).

Perhaps the most productive way to approach the enrollment flow patterns of Figure 3 is to look for the ways academic success probabilities are conditioned by each study phase students pass through, with an eye on practical policy implications. From that angle, the main findings can be stated as follows:

- In dropout probability terms, the first semester does seem to be a particularly dangerous period for PGCC students. The negative exiting tendency of developmental Term-1 students (P0a) fell within the middle of the all-phases range (in fact, it came to less than half the rate exhibited by P0b *developmental* students beyond their first fall). The dropout proclivities of *non*-remedial Term-1 students (P1) registered the second lowest dropout rate of any phase. The prominence of P0a students in Tables 1's negative outcome report (30 percent of all dropouts) was mainly a product of this phase's high participation rate (76 percent).

- Negative outflow ranged from a high of 48 percent in P0b (latter term developmental study) to lows of 10 and 12 percent for P1 (Term 1 non-developmental student) and P4 (degree concentration students). But despite these variations, the best summary of the data is that the "at-risk" problem, to some significant degree, extended across almost the entire academic careers of PGCC students. No phase or phase types monopolized it and it was wholly absent from none. The clear policy implication is that the college needs to broaden its future retention efforts far beyond traditional first term and developmental interventions.

- Type of dropout ("good student" versus "poor student") is also important, since each type of at-risk student requires a different sort of assistance and support. According to Figure 3, early phase negative outflow tended to feature Cohort 1995 students with serious academic problems while more advanced phase dropout was dominated by students with passing course records. The proven substantiality of the phenomenon of the adequately performing student, who abandons his or her academic career just short of graduating or transferring, is particularly distressing. The college has many academic support programs for the academically challenged but offers relatively little structured assistance to academically talented student who may be struggling against ill health, financial problems, and the competing demands of employers and families, the probable root causes of enrollment attrition in this group.

- Figure 3 also provides some useful data on *positive* phase outflow. As is the case at a great many community colleges, so it is at PGCC ? *successful* students tend strongly to transfer *before* graduating. Figure 3 shows that transfer-only likelihood among Cohort 1995 students was a function of the last academic career phase reached. Little such transfer occurred in the first term but, its frequency increased rapidly and monotonically over the post-first-term phase sequence (from P0b 11 percent to a P4 52 percent). In fact, transfer tendency in P3 and P4 was so robust that it became, ironically, the principle cause of failure to graduate. This finding puts a premium on efforts to promote the value of the associate degree, to reinvigorate the transfer degree advisement process, and to improve and increase the number of the college's transfer articulation agreements.

- Finally, as in the dropout case, type of transfer turns out to vary systematically by academic career phase. Although overall transfer outflow was slight in the early phases, what there was of it was of the *lateral* sort. This suggests a "bail-out" motive for shifting enrollment (using transfer to another community college as an alternative to dropping out of higher education altogether). At the other end of student academic careers, four-year transfers predominated. Even so, discernable minorities of transfer exiters at gateway and general education points in their PGCC careers shifted to two-year schools (4 and 3 percent of phase students, and 26 and 12 percent of phase transfers, respectively), which raises the possibility the market forces might be at work.

Conclusions

The purpose of this paper has been to describe a new enrollment flow model, based on the real-world features of a typical community college academic process as experienced by students, to test the validity of that model in terms of fit with longitudinal cohort data, and to demonstrate the practical diagnostic utility of model-based outcomes analysis for enrollment management and program planning. Unfortunately, space limitations have precluded more than a somewhat sketchy presentation of the model's nature and operation, and many of our empirical findings and practical observations had to be set forth in underdeveloped form. We hope, however, that we have managed at least to establish the high potential of our approach for policy-relevant institutional research.

Furthermore, we very much consider our academic career model of student enrollment flow a work-in-progress rather than a finished product, and recognize that much needs to be done to substantiate its method and to improve its explanatory and prescriptive power. We are currently pursuing those ends in a number of ways. First, we are working to refine model component configuration and are struggling to bring

genuine catalog-specified submodeling to the complex degree concentration phase. Secondly, we are moving toward conferring on the model the power to describe and assess the details of within-phase enrollment flow so that college planners can get a more pointed sense of what program reforms are needed to enhance retention, given each phase's internal dynamics. Thirdly, we are exploring model-based enrollment analysis by student body subgroups (full-timers and part-timers; academic, social and financial support program participants; employment status and family status groups; age, gender and race/ethnic groups). Lastly, it seems to us that the academic career flow model, now purely descriptive, could fairly easily be transformed into a predictive instrument by applying event history analysis to phase flows, a possibility we are actively considering. However, whatever the technical progress we might manage to achieve along these lines, what matters most to us is to have succeeded in showing, by means of the present model, however flawed or unfinished, the likely benefits to outcomes research of taking the nuts and bolts of the academic process into systematic account .

References

Andrade, S. (1999). The Heart of the Matter: How Curriculum Affects Student Retention. Paper presented at the Annual Forum of the Association for Institutional Research , Seattle, WA

Astin, A. (1977). Four Critical Years: Effects of College on Beliefs, Altitudes, and Knowledge. San Francisco, CA: Jossey-Bass Publishers, 1977.

Astin, A. (1993). What Matters in College: Four Critical Years Revisited. San Francisco, CA: Jossey-Bass Publishers, 1993.

Bean, J. (1982). Student Attrition, Intentions, and Confidence: Interaction Effects in a Path Model. Research in Higher Education, vol. 17, pp. 291-319.

DesJardins, S., Moye, M. (2000). Studying the Timing of Student Departure from College. Paper presented at the Annual Forum of the Association for Institutional Research, Cincinnati, OH.

Guerin, K (1997). Tracking Student Exit, Transfer and Graduation: An Event History Analysis of Competing Risks. Paper presented at the Annual Forum of the Association for Institutional Research, Lake Buena Vista, FL.

Johnson, R. (1974). Introduction to the Student Flow Model SFM-IA. Technical Report No. 41A. National Center for Higher Education Management Systems [ED111404]

Kelly, L. (1996). Implementing Astin's I-E-O Model in the Study of Student Retention: A Multivariate Time Dependent Approach. Paper presented at the Annual Forum of the Association for Institutional Research, Albuquerque, NM.

Kramer, A., LaMar, A. (2000). Restructuring the General Studies Program at a Public Urban University: Assessment, Evaluation and Implementation. Paper presented at the Annual Forum of the Association for Institutional Research, Cincinnati, OH.

Ronco, S. (1995). How Enrollment Ends: Analyzing the Correlates of Student Graduation, Transfer and Dropout with a Competing Risk Model. Paper presented at the Annual Forum of the Association for Institutional Research, Boston, MA.

Schartman, L., Rhee, B. (2000). Linking Student Retention Model with Institutional Planning: The Benefits and Limitations of a Student Matrix Model. Paper presented at the Annual Forum of the Association for Institutional Research, Cincinnati, OH.

St. John, E., et al. (1991). Effects of Student Financial Aid on Persistence: A Sequential Analysis. Review of Higher Education, vol. 14, no. 3, pp. 383-406.

Tinto, V. (1987). Leaving College: Rethinking the Causes and Cures of Student Attrition. Chicago: University of Chicago Press.

Tinto, V. (1993). Leaving College: Rethinking the Causes and Cures of Student Attrition, 2nd Ed. Chicago: University of Chicago Press.

Young, M. Et al. (1982). Student Flow and Curriculum Matrix, Paper Presented at the Annual Forum of the Association for Institutional Research, May 1982.

PROMOTING A GLOBAL PERSPECTIVE THROUGH INSTITUTIONAL RESEARCH

Anne Marie Delaney
Director of Institutional Research
Babson College

Purpose. The purpose of this paper is to demonstrate how institutional research can promote an enlightened global perspective by enhancing understanding of international students' characteristics, values and goals. The paper is based on a completed research study designed to identify similarities and differences between international and domestic students attending a private college in the northeast section of the Unites States. The ultimate goal of this study was to guide the design of programs that would encourage integration by fostering appreciation of common bonds and respect for differences. The major research questions addressed were:

- First, how do international and domestic freshmen differ with respect to academic quality; interests and activities; perception of abilities; reasons for going to college; and life goals?
- Second, what are the policy and program planning implications of these differences for admissions, marketing, academic management and student affairs?

Review of The Literature. A review of the literature highlights the significant presence, potential influence and challenges associated with increasing numbers of international students attending college in the United States. Foreign students in U.S. regions and states have increased dramatically from 48,486 in l959-l960 to 514,723 in l999-2000. Further, from l998 to l999, foreign student enrollment has increased 4.8 percent nationally. (Chronicle of Higher Education, November 17, 2000).

International students offer a wealth of benefits to our country. As Keith Geiger, Deputy Assistant Secretary for academic programs in the U.S., observes, their presence ensures that there will be a cadre of people around the world who have a profound understanding of the United States which will potentially lead to enhanced relationships and increased cultural ties. (Desruisseaux, l999). There is often, however, a discrepancy between the potential for cultural enrichment and the reality experienced by international students. "On many campuses, often to the dismay of faculty, students from culturally parochial backgrounds quickly link socially with similar others. Even in these globally cosmopolitan contexts, American and international students generally live in separate societies, hardly brushing by each other on sidewalks" (Seymour & Messinger, 1995, p. 4).

Previous research has identified various personal, social, cultural and educational differences as factors that may account for the lack of interaction between international and domestic students

on U.S. college campuses. Tompson and Tompson (l996) reported that international students identified developing a social network, dealing with a different language, and coping with different norms and regulations as the most critical and difficult aspects of adjustment. Faculty cited little class participation, not asking for clarification, and sitting and studying only with international students as behaviors that undermine international students' academic performance.

Three studies identified educational differences that may affect relationships between international and domestic students. Hamilton (l979) reported that international students perceived the university environment to be more competitive, the competition for grades to be more intense, and the professors to be more demanding. Boyer and Sedlacek (l986) discovered that international students have a different perspective on their education and the university environment. They take their education quite seriously, valuing it both for the intrinsic reward of academic pursuit and for career-related reasons. Recently, Ladd and Ruby (l999) found that international MBA students' adjustment involved shifting from the lecture method to a free learning environment in which they had to solve problems rather than memorizing facts. Students needed to be more independent in their approach to learning instead of relying on their professors.

Cultural differences may also affect interaction between international and domestic students. Moline and Hendel (l992) found that, compared with students for whom English was their native language, non-native speaking students put more emphasis on their parents' wishes and the desire to gain a general education as reasons for going to college. They were more interested in being successful in their own business, helping others in difficulty, and promoting racial understanding. In another study (Davis et al., l993), female international students attending an American university identified lack of attention to family, lack of appreciation for the arts, and relatively low interest in reading in the United States as significant cultural differences compared with the values of their country.

Data Source. This study is based on analyses of trend data from the Cooperative Institutional Research Program (CIRP) survey - a national, longitudinal survey of first-time, full-time freshmen in the United States. The survey is sponsored by the American Council on Education (ACE) and the Graduate School of Education and Information Studies at the University of California, Los Angeles. The data set included cases for 1,000 freshmen in the l997, l998 and l999 freshman classes at the study institution.

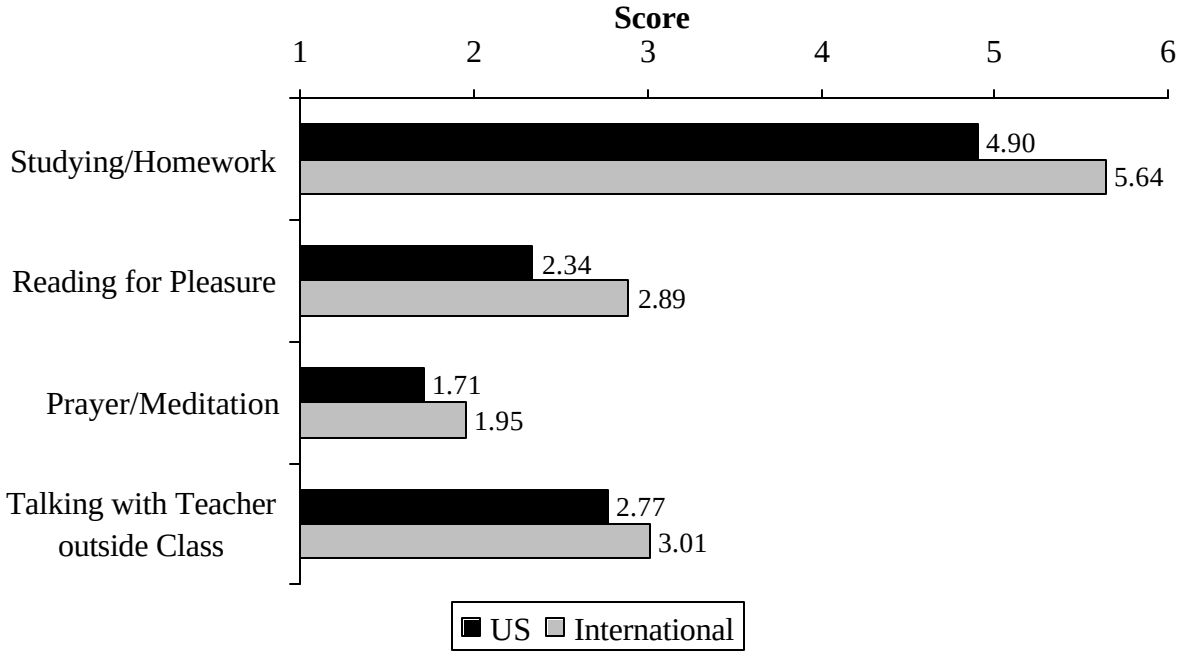
Methods of Analysis. Two-way analyses of variance were used to investigate significant differences by year, citizenship, and the interaction of year and citizenship on: hours spent in various activities during last year of high school, students' academic quality, self-ratings, reasons for going to college, and future goals. Discriminant analysis was employed to identify the most significant differences among international and domestic, male and female students.

Results

Results revealed statistically significant differences by year and citizenship in several areas, including: typical high school activities, academic quality, student self-ratings, reasons for going to college, and life goals.

Differences in Activities and Interests. Results revealed an interesting pattern of differences between international and domestic students in terms of time spent in activities during the last year in high school. As illustrated in Figure 1A, international students spent significantly more time studying and doing homework ($F = 27.74$, $p \leq .001$); reading for pleasure ($F = 21.22$, $p \leq .001$); praying or meditating ($F = 6.97$, $p \leq .01$); and talking with a teacher outside of class ($F = 6.38$, $p \leq .01$).

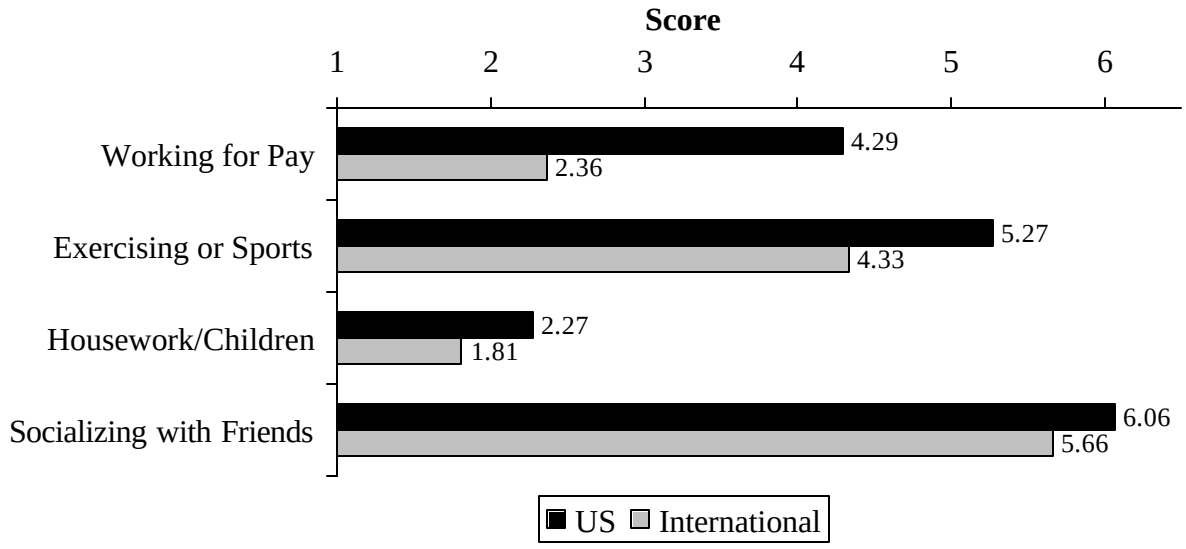
Figure 1A. Significant Differences by Citizenship in Time Spent in Activities During Last Year in High School



Note: Means are based on an 8 point scale from 1 'None' to 8 'Over 20 hours per week'.

In contrast, as shown in Figure 1B, domestic students spent significantly more time working for pay ($F = 68.73$, $p \leq .001$); exercising or playing sports ($F = 41.48$, $p \leq .001$); doing housework or caring for children ($F = 19.03$, $p \leq .001$); and socializing with friends ($F = 10.41$, $p \leq .001$). These differences regarding how students spent their time in high school are important for student programming as they suggest that students will prefer different types of

Figure 1B. Significant Differences by Citizenship in Time Spent in Activities During Last Year in High School



activities during college.

Note: Means are based on an 8 point scale from 1 'None' to 8 'Over 20 hours per week'.

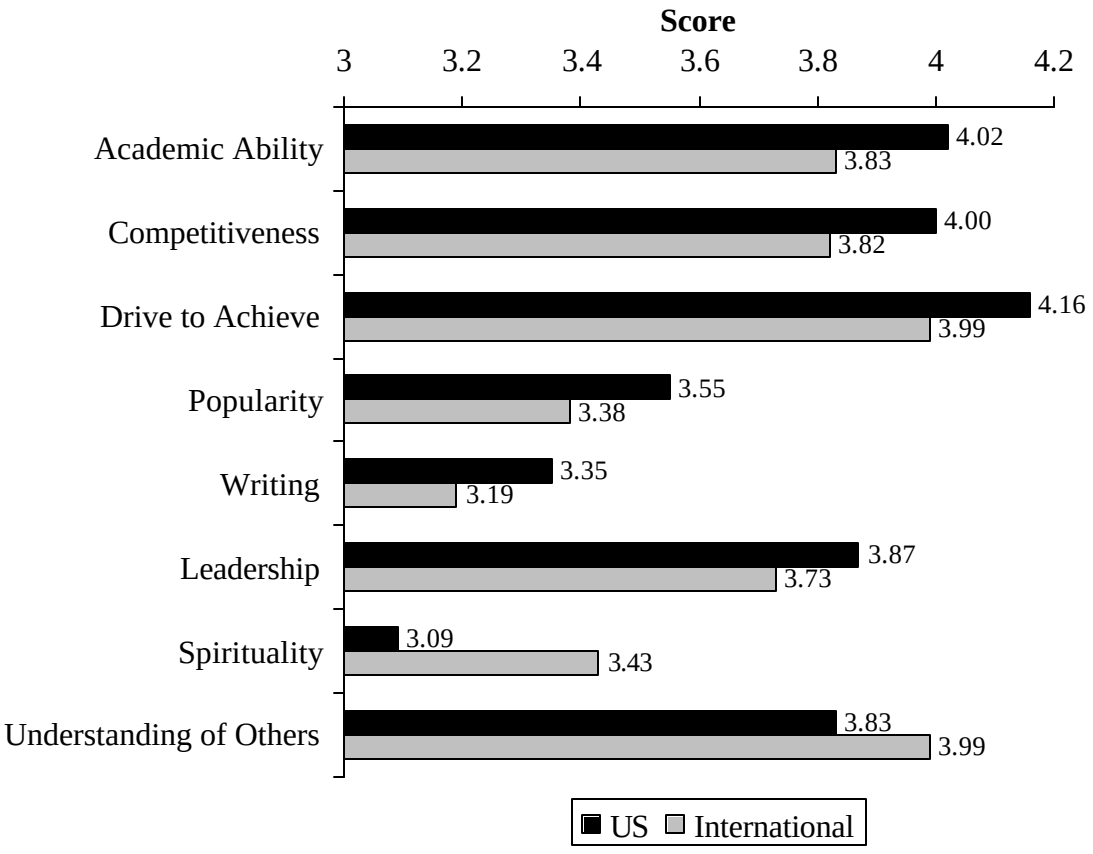
Variation in Academic Quality. Analyses of variance revealed statistically significant differences by year on two measures of academic quality. High school grade point average increased significantly from l997 to l999 ($F= 3.95$, $p \leq .05$). Admission test scores also increased significantly by about 20 points for all students during this three-year period. The SAT Verbal score increased from 569 to 580 ($F = 8.82$, $p \leq .001$) and the SAT Math score increased from 621 to 643 ($F = 4.52$, $p \leq .05$).

With regard to the primary focus of this study, significant differences were found by citizenship status on the SAT Verbal score ($F = 68.44$, $p \leq .001$). The domestic score of 588 significantly surpassed the 531 score for international students. Analyses also revealed significant interaction effects on high school grade point average ($F = 3.54$, $p \leq .05$) and the SAT Verbal score ($F = 4.44$, $p \leq .05$). On high school grade point average, the domestic mean decreased slightly from 6.19 to 6.16 while the international mean increased substantially from 5.67 to 6.41. On the SAT Verbal score, the domestic mean increased somewhat from 579 to 594, while the international SAT verbal mean increased substantially more from 508 to 566. These results indicate that international students' academic quality has improved at a faster rate than that of domestic students.

Students' Self Ratings of their own Abilities. Both international and domestic students' self-ratings increased significantly by year on the following four characteristics related to self-confidence and specific academic abilities: mathematical ability ($F = 4.07$, $p \leq .05$), intellectual self-confidence ($F = 4.82$, $p < .01$), social self-confidence ($F = 3.95$, $p \leq .05$), and writing ability ($F = 3.46$, $p \leq .05$).

Analysis by citizenship revealed that domestic students rated themselves significantly higher on several characteristics and abilities including: academic ability ($F = 12.58$, $p \leq .001$), competitiveness ($F = 14.12$, $p \leq .001$) and drive to achieve ($F = 7.59$, $p \leq .01$). In contrast, international students rated themselves higher on spirituality ($F = 13.14$, $p \leq .001$) and understanding of others ($F = 6.49$, $p \leq .05$).

Figure 2
Significant Differences by Citizenship in Students' Self Ratings

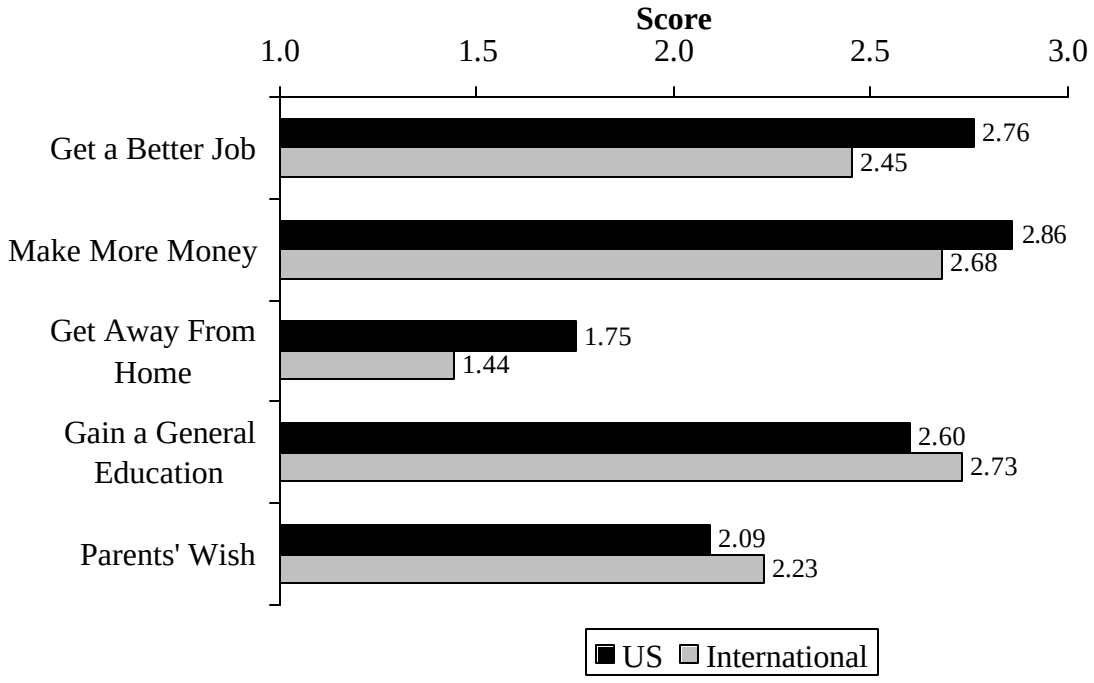


Note: Means are based on a 5 point scale from 1 'Lowest 10%' to 5 'Top 10%'.

How students perceive themselves bears important implications for curriculum development and student program planning. These data identify perceived strengths and weaknesses and indicate areas that should be addressed in educational and student affairs programs.

Variance in Reasons for Going to College. International and domestic students differed significantly on reasons for going to college. As illustrated in Figure 4, domestic students attributed more importance to getting a better job ($F = 31.05$, $p \leq .001$) and making more money ($F = 21.00$, $p \leq .001$). In contrast, international students placed more importance on gaining a general education ($F = 7.76$, $p \leq .01$) and their parents' wishes ($F = 5.12$, $p \leq .05$).

Figure 3. Significant Differences by Citizenship in Reasons for Going to College

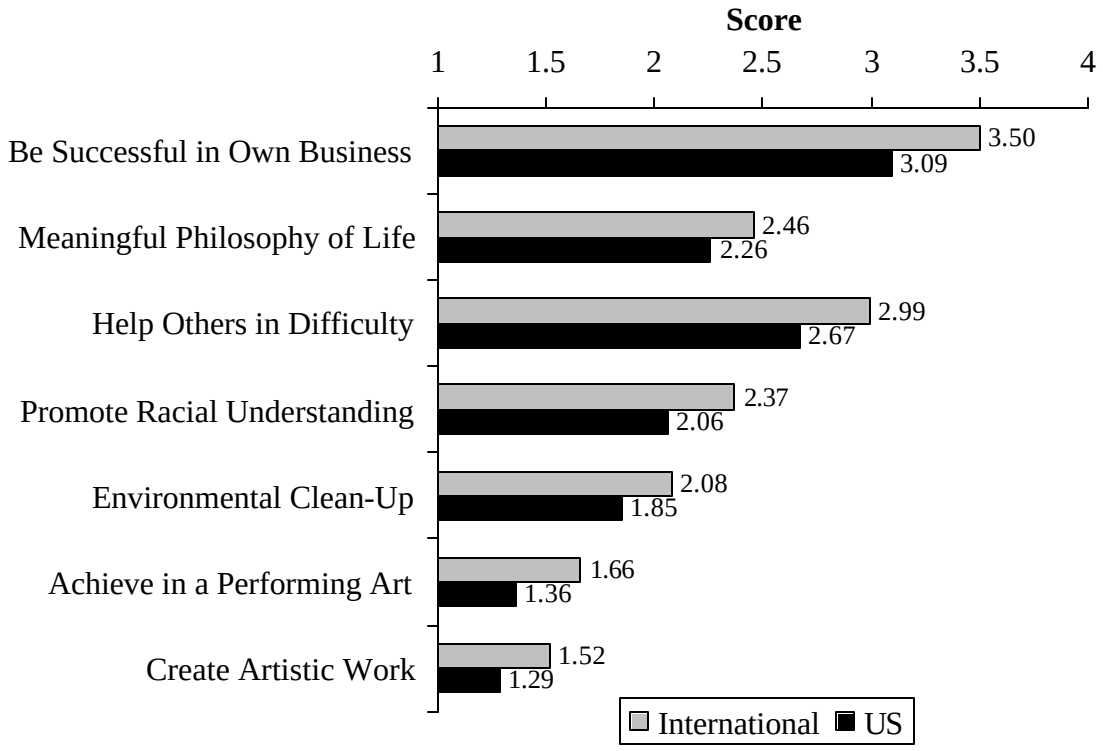


Note: Means are based on a 3 point scale from 1 'Not Important' to 3 'Very Important'.

These differences regarding reasons for going to college reflect student values and suggest that students will have different expectations from college. This information is important both for admission recruiting and undergraduate program planning. One implication is that an undergraduate program that offers a general education, as well as excellent career preparation, will be attractive to international students.

Significant Variation in Life Goals. Analysis by citizenship revealed that international students differ significantly from domestic students on several future goals. As displayed in Figure 4, international students attribute more importance to being successful in their own business ($F = 28.94$, $p \leq .001$), developing a meaningful philosophy of life ($F = 5.79$, $p \leq .05$), helping others in difficulty ($F = 23.27$, $p \leq .001$), and promoting racial understanding ($F = 17.86$, $p \leq .001$). Some of these life goals are consistent with international students' emphasis on going to college to gain a general education and may indicate different expectations from the college experience as a preparation for life.

Figure 4. Significant Differences by Citizenship in Goals



Note: Means are based on a 4 point scale from 1 'Not Important' to 4 'Essential'.

Results from Discriminant Analysis. To provide more information, the analysis was expanded to include gender as well as citizenship status. Discriminant analysis was used to predict membership in one of the following four groups: male U.S. citizens, female U.S. citizens, male international students, and female international students. Two statistically significant functions were identified.

Table 1A displays the means for each of the four groups on function 1. Table 1B identifies several variables correlated with function 1. As shown, function 1 contrasts male U.S. citizens with U.S. females and international males and females. A review of the group means on this function shows that U.S. males are most dissimilar to international females. It is also interesting to note that U.S. females are more similar to international students than to U.S. males. The statistical results on function 1 indicate that U.S. male students are typically high ability, confident, competitive students who share the desire to make more money as an important reason for attending college. In contrast, U.S. females and international students share more spiritual values. They are more interested in helping others in difficulty, in promoting racial understanding, and in achieving in a performing art. Further, parents' wishes and a desire to gain a general education are important reasons for U.S. female and international students' decision to attend college.

Table 1A. Summary of Canonical Discriminant Function

Function 1	**Group Centroids**
Male U.S.	+.50
Male International	-.51
Female U.S.	-.42
Female International	-1.32
Chi-Square = 394.66; df = 84; $p \leq .001$	

Table 1B. Structure Correlation Coefficients with Function 1

Variables	**Correlation Coefficient**
SAT Verbal Score	.52
Competitiveness	.47
Leadership Ability	.27
Intellectual Self-Confidence	.29
Math Ability	.27
Academic Ability	.27
SAT Math Score	.25
Desire to Make More Money	.21
Emotional Health	.21
Spirituality	-.22
Reason for College– Parents' Wish	-.25
Reason – Gain a General Education	-.27
Help Others in Difficulty	-.28
Promote Racial Understanding	-.29

Achieve in a Performing Art	-.30

Function 2 contrasts international students with U.S. citizens, particularly U.S. male students. Table 2A displays the means for each of the four groups on function 2. Table 2B identifies several variables correlated with this function. According to these data, international students perceive themselves as more spiritual and creative. They are more confident intellectually and are more interested in the goal of being successful in their own business. Also, U.S. male students, compared particularly with international male students, have higher SAT verbal scores, higher average high school grades, and are more focused on attending college to get a better job.

Table 2A. Summary of Canonical Discriminant Function

Function 2	**Group Centroids**
Male International	+.97
Female International	+.49
Female U.S.	.009
Male U.S	-.65
Chi-Square = 170.14; df = 54; $p \leq .001$	

Table 2B. Structure Correlation Coefficients with Function 2

Variables	**Correlation Coefficient**
Be successful in own business	.48
SAT math scores	.38
Intellectual self-confidence	.23
Creativity	.23
Self-understanding	.21
Achieve in a performing art	.18
Spirituality	.11
Reason for College – Get a better job	-.23
SAT verbal scores	-.26
Average high school grades	-.35

As shown in Table 3, the discriminant analysis accurately predicted 64 percent of all cases. By far, the highest prediction rate is for U.S. male students. In contrast, the prediction rate is quite low for international students, particularly male students.

Table 3. Classification Results: Predicting Group Membership

	Percent Correctly Classified
Male U.S.	86.8%

Female U.S.	43.9
Male International	16.9
Female International	32.3
All	64.1%

Results from this analysis suggest the need for more information to better understand international and domestic female students. Further research should seek to identify their unique qualities, values, aspirations, and goals. Insight gained from such research will potentially enable the College to better serve these students and to promote greater interaction among international and domestic male and female students on campus.

Discussion

Results from this research advance our understanding of significant differences between international and domestic students. The following discussion relates findings from this study to results from previous research. Knowledge from this body of research provides a basis for planning programs to promote enhanced relationships and to realize the potential benefits of international students' presence at colleges in the United States.

In their study on international students' adjustment, Tompson and Tompson (l996) found that developing a social network and coping with different norms were among the most difficult aspects of adjustment. Results from this study suggest that differences in interests, goals and values may account for some of the adjustment difficulties. Compared with domestic students, international students spent more time reading, studying, praying or meditating and conversing with their teachers; whereas domestic students spent significantly more time working for pay, exercising or playing sports, and socializing with friends. These differences are important as they suggest that students will prefer different types of activities during college.

Similar to findings reported by Moline and Hendel (l992) and Davis, et al. (l993), international students in this study also expressed greater interest in various social, philosophical and artistic goals, including developing a meaningful philosophy of life, helping others in difficulty, and achieving in a performing art. Recognizing and understanding these differences in values is essential to developing programs that will be of interest to international students and to fostering meaningful relationships between international and domestic students.

How students perceive themselves also affects relationships. Analysis by citizenship revealed that domestic students rated themselves significantly higher on several characteristics and abilities, including: competitiveness, drive to achieve, and academic ability. In contrast, international students rated themselves significantly higher on spirituality and understanding of others.

Similar to results from previous studies (Boyer & Sedlacek, l986; Moline & Hendel, l992),

findings from this research also show that international and domestic students have different reasons for going to college. Domestic students attributed more importance to getting a better job and making more money. In comparison, international students placed more importance on gaining a general education and their parents' wishes.

Results from this study were translated into strategic policy recommendations for the President's Cabinet, the Undergraduate Admission, Academic Program Management, and Student Affairs Departments at the study institution. The overall goal of these recommendations was to enlighten future policy and guide the design of programs that would encourage integration by fostering appreciation of common bonds and respect for differences. Illustrative recommendations follow.

Recommendations

1. **In recruiting international students, emphasize the College's ranking in entrepreneurship.**

These data indicate that international students are significantly more interested than domestic students in being successful in their own business.

2. **Design undergraduate admission publications and communications to reflect international students' intellectual, cultural and social values.**

Results from this study reveal that international students spent significantly more time studying, reading for pleasure, and praying or meditating during high school, and they report significantly higher interest in certain personal and social goals: developing a meaningful philosophy of life, helping others in difficulty, and promoting racial understanding.

3. **Emphasize opportunities for students to communicate with the faculty.**

Compared with domestic students, international students reported they spent significantly more time talking with a teacher outside of class in high school. Such experience would potentially lead them to expect similar opportunities during college.

4. **Portray the College experience as an opportunity to gain a general education as well as to prepare for a career.**

In discussing their reasons for going to college, international students placed significantly more importance on gaining a general education while domestic students attributed more importance to getting a better job and making more money.

5. **Design programs to identify and serve students in need of language support.**

Overall, international students' SAT verbal scores were significantly lower than those of domestic students. While the international student mean score increased from l997 to l999, it still remained lower than that of domestic students.

6. **Offer seminars and discussion groups on culture and values for international and domestic students to learn about each other's cultural values.**

Results from this study reveal that international and domestic students differ significantly in their reasons for going to college and in their interest in pursuing many personal and social goals. Some of these differences may reflect differences in cultural values. Offering students an opportunity to learn about different cultures may foster greater understanding and ultimately promote increased interaction between international and domestic students.

7. **Develop social programs, highlighting different international cultures, to ensure that international students feel welcome and domestic students learn to appreciate the values and customs of international students' countries.**

Findings from national studies reveal that American and international students often live in separate societies while attending the same college. The opportunity for cultural enrichment associated with international students' presence is often missed.

8. **Evaluate the undergraduate program to ensure that it provides international students with opportunities to achieve growth in a broad range of life goals.**

Comparative analyses identified statistically significant differences between international and domestic students on several goals, including: developing a meaningful philosophy of life, helping others in difficulty, and promoting racial understanding. These life goals are consistent with international students' emphasis on going to college to gain a general education and may indicate different expectations from the college experience.

9. **Undergraduate program and course planning need to take into account the increase in student quality from l997 to l999. Special attention should be given to ensure that academic courses and programs keep pace with student quality.**

From l997 to l999, admission test scores increased significantly among entering freshman classes. Scores increased by about 20 points from 569 to 589 on the SAT Verbal score and from 621 to 643 on the SAT Math score.

10. **The College should continue to monitor international student trends and examine the extent to which international and domestic students interact and enrich each**

other's lives.

While the presence of international students offers the potential for cultural enrichment, experience suggests that the College may need to exercise initiative to realize this potential.

References

Boyer, S. and Sedlacek, W. (1986). Attitudes and perceptions of incoming international students (Report No. RR-4-86). College Park, MD: Maryland University. (ERIC Document Reproduction Service No. ED 278 935)

Davis, W. and others. (1993). Cultural practices of female international students. Austin, TX: University of Texas Austin. (ERIC Document Reproduction Service No. ED 356 739)

Desruisseaux, P. (1999, December 10). Foreign students continue to flock to the U.S.. Annual report finds that economic turmoil in Asia did not lead to a decrease in enrollments. The Chronicle of Higher Education, A57-A58.

"Foreign students in U.S. regions and states, selected years, 1959-60 to 1999-2000." (2000, November 17). The Chronicle of Higher Education, [On-line]. Available: http://chronicle.com/weekly/v47/i12/12a07605.htm

Hamilton, J. (1979). A comparison of domestic and International Students' Perceptions of the University Environment. Journal of College Student Personnel, 443-445.

Ladd, P. and Ruby, R. (1999). Learning style and adjustment issues of international students. Journal of Education for Business, 74 (6), 363-367.

Moline, A. and Hendel, D. (1992, May). Exploring international issues through the use of CIRP data. Paper presented at the annual forum of the Association for Institutional Research, Atlanta, GA. (ERIC Document Reproduction Service No. ED 349 864)

Seymour, R. and Messinger, S. (1995). Inviting international students into the communication classroom: The Wayne State University global communication course. Detroit, MI: Wayne State University. (ERIC Document Reproduction Service No. ED 389 013)

Tompson, H. and Tompson, G. (1996). Confronting diversity issues in the classroom with strategies to improve satisfaction and retention of international students. Journal of Education for Business, 72 (1), 53-57.

THE ROLE OF INSTITUTIONAL RESEARCH IN IMPLEMENTING FIRST-YEAR SEMINARS

Michael J. Dooris
Director, Planning Research and Assessment

Daniel P. Nugent
Management Information Associate

The Pennsylvania State University

Abstract

In 1999-2000, Penn State implemented a first-year seminar requirement for all incoming baccalaureate freshmen, in all majors, at all Penn State campuses. This paper describes the role that institutional research played through three stages of the process:

- *early committee deliberations;*
- *implementation of the new program;*
- *support and development of the program.*

Student and faculty focus groups and surveys have provided mostly positive feedback about the seminars, along with some insights into opportunities for refinement and improvement. Analyses of non-obtrusive data such as transcript files and registrar's databases have been an especially efficient and powerful way to answer questions about how the new requirement has been enacted – To what extent are students fulfilling the requirement? Through which courses? For how many credits? In which colleges? – and so on.

At Penn State, we have emphasized relatively pragmatic analyses (such as how the curriculum is being enacted) more than classic assessment approaches (such as pre- and post-tests), for two reasons. First, the challenges of conducting valid and reliable assessments of cognitive and affective gains, or educational outcomes, are substantial, especially for campus IR staffs juggling multiple responsibilities with limited resources of time, staffing, money, and expertise. Second, a persuasive research literature has demonstrated that first-year seminars do constitute good practice. Therefore, our own efforts have mostly been directed toward developing, supporting, and strengthening a program that faculty and administrators believe is a good idea. We suspect that these probably are legitimate considerations for many of our colleagues in institutional research offices at other campuses, as well.

Introduction

Beginning in 1999-2000, Penn State University initiated one of the nation's most ambitious first-year seminar efforts when it enacted a university-wide first-year seminar requirement for all new incoming baccalaureate students. Institutional research was involved in all phases of the design, implementation, and assessment of the first-year seminar program. This paper draws from that experience to illustrate how institutional researchers can contribute to this

type of curricular change.

Origins of the First-Year Seminar Program at Penn State

The first-year seminar program at Penn State was part of a larger change in general education at the university. In December 1997, Penn State's Faculty Senate adopted the recommendations of a special committee, which had been working for over a year to develop a new general education curriculum (Penn State, Fall 1997).

Some of the changes in the general education package were more substantial and challenging than others, but in total, the adjustments were significant. The committee presented ten recommendations (all of which the Faculty Senate ultimately accepted) with the goal of enhancing curricular flexibility, emphasizing high quality, fostering opportunities for experimentation and building assessment into the curricular process. Among the committee's recommendations were the inclusion of active learning elements in all general education courses and the identification of key competencies. The committee also recommended the restructuring of requirements for health and physical education and for foreign languages. (More detailed information is available at Penn State's general education website, http://www.psu.edu/oue/gened/.) The hallmark of the new general education curriculum was the establishment of a first-year seminar requirement for all incoming baccalaureate students.

Members of the general education committee certainly drew on their own experiences and ideas in developing their report, but more objective information from qualitative and quantitative analyses was also important at all stages of the process: in initial explorations into alternatives and possibilities, through the implementation of curricular changes, to the assessment of the new program (Dooris and Blood, 2001).

The Role of Institutional Research

Institutional research supported the committee with data from many sources throughout the three stages of early deliberation, program implementation, and program support and continuation.

Early Deliberations about a First-Year Seminar Program

Early in the discussion process, institutional research helped confirm the desirability of some type of program along the lines of a first-year seminar. For example, committee members learned from transcript analyses that 55 percent of general education student credit hours were earned in classes of 100 or more students and that freshmen and sophomores were disproportionately likely to be enrolled in large classes. Attendance studies revealed that student absenteeism was strongly and positively correlated to class size; alumni and student surveys highlighted a need for students to better see the linkages between general education with studies in the major, especially early in the college experience. Several colleges within the Penn State system were already using elective first-year seminars and data were available from the assessment of these existing programs. The literature on undergraduate education and on first-

year seminars as well as benchmark information on practices at other colleges (not all of which were in Penn State's normal research university peer group) was also helpful.

First-Year Seminar Implementation

Institutional research also played a very useful role during the implementation stage. The actual creation of a first-year seminar program is a significant undertaking for any institution; it is especially ambitious for an institution as large and organizationally complex as Penn State. Each year Penn State enrolls about 12,000 new first-year students on its 24 campuses. It has 17 undergraduate degree-granting colleges that offer majors in 232 undergraduate programs. Every college or university is, of course, unique – but institutional research can help faculty, staff, and administrators at any institution decide whether and how a major curricular change can, as a practical matter, be designed and implemented in the context of practical and specific organizational considerations.

In Penn State's situation, the general education committee was interested in proposing a university-wide requirement, but had doubts about whether the university would be able to overcome some very significant obstacles. Institutional research helped address these doubts by developing information on factors such as existing patterns of faculty instructional assignment; curricular requirements (credit loads, course sequencing) of different majors; estimated numbers and costs of additional small sections; and physical facility constraints—that is, the number, size, and availability of classrooms.

Taking such factors into consideration, the committee decided that an institution-wide program could work if it combined a university-wide philosophy with considerable college, campus, and departmental flexibility.

University-Wide Approach

The most important university-wide aspect is the fact that there is a Penn State first-year seminar requirement that applies to all entering baccalaureate freshmen, in every academic program of the university and at every campus. We believe Penn State is the only large research university in the nation with such an institution-wide provision. Also, the seminars are all credit-bearing, all taught by regular full-time faculty with at least three years teaching experience at Penn State, and all limited to twenty students per section.

College and Department Flexibility

While the program has strong university-wide common threads, the first-year seminar is nonetheless as varied and individual as Penn State's many colleges, campuses, and departments.

For example, the first-year seminar courses range from one to three credits. Most students take a seminar in the college in which they are enrolled, but they may select from a wide array of courses, and satisfaction of the requirement is portable. Colleges, departments and campuses design and offer their own courses and have considerable autonomy as to the content and structure of the offerings. In recognition of the difficulty in implementing a university-wide program to serve 12,000 students, the Faculty Senate and the University administration specified a two-year

transition period between acceptance of the recommendations in 1997, and enactment of the first-year seminar requirement for freshmen entering in the Summer and Fall of 1999.

Assessment of the Program

Institutional researchers have been and remain involved in assessment of the first-year program, and in reporting results annually to the Faculty Senate and to the deans' council.

Consistent with the idea that assessment should be designed into a new program, one of the ten recommendations of the general education committee in fact called for systematic assessment of general education. As a result, the Faculty Senate and the university administration appointed a nine-member General Education Assessment Interest Group. That group identified the assessment of first-year seminars as its first task; it helped to guide much of the institutional research described in this paper.

The Need for Footprint Assessment

First-year seminar programs are frequently asked to prove their value; this theme threads throughout the publications of South Carolina's highly regarded National Resource Center for the First-Year Experience.

There is often a temptation to focus such assessments on the most conceptually interesting research questions about gains, cognitive and affective outcomes, and the like, but every college or university should probably first ask, "To what extent is this institution actually enacting its supposed curricular requirements? What is the *footprint* of programs, in terms of offerings and course-taking patterns?" Information about course offerings and student enrollment patterns are prerequisite to an evaluation of the strengths and weaknesses of a program. This is particularly true in the case of Penn State, given the myriad of first-year seminar implementation models among the individual colleges and campuses. Fortunately for institutional researchers, data about these patterns are available unobtrusively from transcript data, registrar's files, and the like. Increasingly, at many colleges and universities, such data can be accessed fairly easily from a central data warehouse.

The Penn State Data Warehouse

In 1994, Penn State began development of a university-wide data warehouse. The aim of the data warehouse was to simplify *ad hoc* access to the most widely used administrative data.

The data warehouse has since grown to approximately 100 tables in more than a dozen databases housing just under 100 million records. Data are transferred from the legacy systems on a regular schedule. These data are non-modifiable and represent a snapshot of time-fixed data. The Penn State data warehouse provides a convenient and consistent source of institutional data and allows for *ad hoc* inquiries as well as extensive analysis.

Using the data warehouse – specifically the data available in the university's official

enrollment, term course master, and transcript files – we were able to describe how the first-year seminar requirement had been enacted. For example, by the end of the 1999-2000 academic year, 93 percent of the first-time enrolled freshmen who returned as sophomores had completed the first-year seminar requirement. There were 764 sections of 234 different courses offered to satisfy the requirement. The sizes of the sections ranged from 14 to 25, with most sections within the desired maximum size of 20 students. The large majority (81 percent) of the courses were offered for one credit, eight percent of the courses were offered for two credits, and 11 percent of the courses were offered for three credits.

Concerns over the university's ability to implement such a significant curriculum change were eased, since these data also showed that more than enough first-year seminar places were available to serve the needs of the student population. The data also have helped to guide the development of procedures to handle situations in which a student does not satisfy the requirement in his or her freshman year.

Student and Faculty Views on the First-Year Seminar Program

Penn State relied largely on faculty and student focus groups and student surveys in gauging the impact of the first-year seminar program and in identifying areas where improvement is needed.

Faculty Focus Group Results. Faculty focus groups representing ten colleges and four campus locations provided feedback that was mostly positive. A strong message was that it is desirable to allow faculty to be creative with the content and structure of their individual first-year seminars while still ensuring that common objectives are addressed. Almost all the faculty members expressed a desire to teach a first-year seminar again.

Student Focus Group and Survey Results. Focus groups were conducted with students from six colleges and four locations, and approximately 500 freshmen who had completed or were currently enrolled in first-year seminars were surveyed via written instruments. As with the faculty feedback, student reactions were largely positive. Students especially liked the small class sizes, which created opportunities for interaction both with faculty and with other students, many of whom were also enrolled in their other courses. Students identified time management skills, academic content, career knowledge of major and field and enhancement of library and Internet or computer skills as the most important things they learned in the first-year seminar.

Surveys also found that almost half of the respondents felt that their first-year seminar resulted in their becoming engaged in the climate of learning at Penn State, being oriented to high expectations and demanding workload of academic life, and seeing a connection between the first-year seminar and their potential major. 35 percent of survey respondents felt that interacting with their first-year seminar instructor "added to the quality of their first year experiences" (Penn State, 2001).

While there were no strong negative comments, student feedback did indicate some

areas for improvement in the program. Students who had completed a compressed seminar reported that too much material was covered in too short a time. Students who had not taken the seminar in their first semester felt that they should have take it then. Students felt that they should have been informed earlier about the array of first-year seminar choices available to them across the university.

Conceptual, Methodological, and Practical Considerations

Higher education now has about two decades of substantive, large-scale experience with first-year seminars, and with the assessment of those programs. Because this paper focuses on the assessment aspects of Penn State's program, it is worth at least briefly discussing what the university did – and did not do – in the context of that broader, national experience.

Penn State's basic approach was essentially pragmatic. The faculty and administration strongly believed in the value of first-year seminars. They also felt assessment was important – but primarily to help create, implement, and continually improve a program that is fundamentally a good idea (versus, for example, to contribute to the scholarly literature on the topic, or to explore a subject of mostly academic interest). In this respect, Penn State was not very different from most colleges and universities. John Gardner, for example, has observed that the freshmen seminar movement has mostly been directed toward developing programs that work, and only secondarily to developing elaborate evaluations of those programs (Gardner, 2001).

We believe that this is a very sensible bias for individual institutions, because the broad picture that has emerged from two decades of work is quite clear.

First-Year Seminars as Good Practice

First-year seminars constitute good practice. A compendium of studies from 50 colleges and universities provides evidence from institutions of all sizes, missions, and selectivity: "Retention rates improve, grades improve, students' internal locus of control increases, participation in extracurricular activities and the use of campus services both increase, and students begin to clarify their short- and long-term goals. Most importantly, graduation rates increase" (Barefoot, 1998, p. xi).

Reviews of numerous studies have shown that the first six weeks of the freshman year are the critical determinant of ultimate graduation (Erickson and Strommer, 1991). The evidence also shows that first-year seminars are an effective method for initiating students to higher education, helping them to make a successful transition from high school (Leamnson, 1999). Upcraft and Gardner's (1989, pp. 4-11) review found that freshmen seminars enrich opportunities for student involvement – vital to freshman success – and that there is "conclusive evidence…that the freshman seminar is a very powerful way of enhancing freshman success."

The Measurement Challenges

Looking at the issue of evaluating excellence in undergraduate education, Ernest Pascarella (2001) recently reviewed some of the significant methodological challenges to validly and reliably measuring quality, excellence, or success. How would an institution decide what outcomes to measure? What particular set of competencies, activities, and accomplishments can be attributed to the undergraduate experience? How does a study control for out-of-class or out-of-college experiences? How does a study control for differences in student ability? And so on.

Our strong suspicion is that few campus institutional research practitioners – while juggling multiple responsibilities with limited resources – have the time, money, staff, and psychometric expertise to overcome these research design challenges on a realistic schedule.

Emphasize Good Practice. Bluntly put, at some level choices must be made between the sort of pragmatic implementation-oriented assessment that Penn State has emphasized, and more classic assessment approaches: pre-test/post-test, quasi-experimental designs, outcome measurements, and so on.

Interestingly – because he was not writing about the assessment of first-year seminars or about institutional research – Pascarella did suggest an approach which is very consistent with most of Penn State's institutional research on first-year seminars. In brief, Pascarella suggested a focus upon the practices and processes that are known to be linked to important cognitive and noncognitive outcomes. He wrote, "the assumption here, and it is not an unreasonable one, is that an excellent undergraduate education is most likely to occur at those colleges and universities that maximize good practices" (Pascarella, 2001, p. 22).

John Gardner has noted that first-year seminars are "the most studied and assessed course genre in American higher education history" (1998, p. xiii). It is important for institutional researchers to take this observation seriously, and to help their respective campuses take advantage of the collective wisdom that higher education has accumulated.

In short, first-year seminars are a good idea. Peter Ewell suggested (2001) that researchers continually ask themselves two concrete, core questions: "What happened?" and "What mattered?" We believe that is useful, wise, and legitimate advice for practitioners involved in institutional research on first-year seminars.

References

Barefoot, B.O. (1998) (Ed.) *Exploring the evidence: Reporting outcomes of first-year seminars*. (Monograph No. 25). Columbia, SC: University of South Carolina, National Resource Center for the First-Year Experience and Students in Transition.

Dooris, M.J., and Blood, I.M. (July-August 2001). Implementing and assessing first-year seminars. *Assessment update, 13* (4), 1-12.

Erickson, B.L. and Strommer, D.W. (1991) *Teaching college freshmen*, Jossey-Bass.

Ewell, P. (2001). Observations on assessing the first-year experience. In Barefoot, B.O. (Ed.) *Exploring the evidence: Reporting outcomes of first-year seminars*. (Monograph No. 25). Columbia, SC: University of South Carolina, National Resource Center for the First-Year Experience and Students in Transition.

Gardner, J.N. (1998). Introduction. In Barefoot, B.O. (Ed.) *Exploring the evidence: Reporting outcomes of first-year seminars*. (Monograph No. 25). Columbia, SC: University of South Carolina, National Resource Center for the First-Year Experience and Students in Transition.

Gardner, J.N. (2001). Preface. In Swing, R.L. (Ed.), *Proving and improving: Strategies for assessing the first college year* (Monograph No. 33). Columbia, SC: University of South Carolina, National Resource Center for the First-Year Experience and Students in Transition.

Leamnson, R. (1999). *Thinking about teaching and learning: Developing habits of learning with first-year college and university students.* Stylus Publishing.

Pascarella, E.T. (May/June 2001). Identifying excellence in undergraduate education: Are we even close? *Change*, 19-23.

Penn State (Fall 1997). *Final report and recommendations of the special committee on general education for the University Faculty Senate.* University Park, PA. (http://www.psu.edu/ufs/gened.html)

Penn State (March 2001). Penn State Pulse, First-Year Learning Outcomes. (http://www.sa.psu.edu/sara/pulse/shtml)

Upcraft, M.L. and Gardner, J.N. (1989). A comprehensive approach to enhancing freshman success. In *The freshman year experience*, Jossey-Bass.

ORGANIZATIONAL AND TECHNICAL ISSUES IN IMPLEMENTING WEB-BASED SURVEYS: A CASE STUDY

Gordon J. Hewitt
Dawn Geronimo Terkla
Heather S. Roscoe
Jennifer Dyer

Tufts University

Introduction

The use of the web in conducting survey research in institutional research is on the rise. The benefits to be gained are numerous: cost savings due to the elimination of copying, mailing, and data entry; "instant data," which allows for quicker turn-around time; and easier access to a wider population, such as international respondents. While the verdict is still out on methodological issues such as response rates and biases, web-based surveys are the trend of the future.

However, there is uncertainty among IR professionals on how to proceed in implementing web-based surveys. The questions are myriad: Do we build our own software or purchase an existing product? Do we need our own server, use a server on campus, or contract to use a private server? Do we need technical staff members in the IR office or can we do it ourselves? The list goes on. One thing is certain. Unless your IR office already has an IT or systems expert, you need to take the time to learn about the different options and what will work best based on the level of technical support, the complexity of the technical infrastructure at your campus, and your own level of technical expertise.

Objective

The primary objective of this paper is to describe the process of one IR office's attempt to select web survey software, coordinate the use of on-campus servers and technical staff, and evaluate the effectiveness of the entire process. This case study will inform and educate other institutional research professionals considering a move from paper surveys to web-based surveys.

Literature

The body of literature on the methodological implications of web-based surveys is growing (Matz, 1999; Cartwright, Thompson, Poole and Kester, 1999; Underwood, Kim and Matier, 2000; Porter and Umbach, 2000). These and other studies discuss at great length the methodological costs and benefits of moving from paper to the web, and greatly inform the IR professional's decision to make that move. The benefits, of course, are clear. It is much cheaper to conduct a web-based survey since postage, printing, and other distribution costs are

not incurred. It is also much quicker to conduct a web-based survey, since data is entered directly by the respondent to an electronic data file with out the need of second-party data entry. The costs of web surveying, however, may have an impact on survey results. Several institutional research projects in higher education reported lower response rates for web-based surveys as compared to paper surveys (Cartwright, Thompson, Poole and Kester, 1999; Underwood, Kim and Matier, 2000). Underwood et *al* also reported that underrepresented minority students responded on the web at a lower rate than White and Asian students. However, some non-higher education studies found no significant differences between paper and web survey responses (Smith 1997).

Currently, there are no studies that specifically address the decision-making process required for choosing the tools for web-based surveys and for dealing with campus technical-organizational issues. While some studies touch on logistical issues, such as Porter and Umbach (2000) and Cartwright, Thompson, Poole and Kester's (1999) description of assigning students unique PIN numbers, there are no comprehensive descriptions of the decision-making process in regards to applying available technology to an IR office's surveying needs.

Surveying Background and History

Description of Structures and Processes

This case study describes the decision-making processes involved in implementing a web-based survey system by the Office of Institutional Research (OIR) at Tufts University. Tufts is a Doctoral/Research University-Extensive located just outside of Boston in Medford, Massachusetts. The university comprises eight schools: Arts & Science, Engineering; Fletcher School of Law and Diplomacy; Nutrition Science and Policy; Medical; Sackler School of Biomedical Sciences; Dental; and Veterinary. The Office of Institutional Research is a centrally-located unit, serving all of the schools of the university as well as the central administration.

The OIR at Tufts is known across campus – to both administrative and academic units -- as a good resource for conducting survey research. Assistance to units ranges from consultation on survey item development to management of entire projects. In fiscal year 2000, the time period immediately before web surveys were used by the OIR on a regular basis, the office administered 24 different surveys. These surveys were both annual projects, such as the senior survey and the accepted applicant surveys, and ad hoc or intermittent projects, such as the undergraduate alumni survey. They also ranged in size from a single page with an N of less than 100 to multi-page instruments with N's over one thousand. The OIR at Tufts has a history of creating their own "homegrown" survey instruments rather than relying on standard national surveys. For example, the OIR, in collaboration with admissions staff, developed and administered a New Student Survey and A Non-Enrolling Survey instead of using the standard Admitted Student Questionnaire survey, which many other college campuses use. The theory behind the development of these homegrown instruments is that it gives the OIR's client – the Office of Admissions – the ability to ask university and issue-specific questions to a greater degree and allows the OIR to have more control over data formatting.

The processes in the OIR were somewhat standard for administering paper surveys with large populations, and return procedures varied based on the project. For example, business reply envelopes were provided for surveys that went off campus to alumni. Surveys on campus, to student or faculty populations, were either returned through campus mail or collected directly from the respondent. Once the completed surveys were returned, the data was entered manually, usually by student research assistants. The office experimented for a period of time with two different scanning procedures, but found that they were not efficient and the time and labor cost differentials compared to manual data entry were not great.

The result of the OIR's success in cultivating a sense of trust and understanding in the importance of data collection to inform and improve programs among its campus clients was a dramatic increase the number and scope of surveys administered. This has been especially true since the mid-to-late 1990's, when the evaluation and assessment movement took hold in higher education Along with this increase, however, came the need for resources to complete these research studies. The cost of most of these surveys ran anywhere from a few hundred dollars to several thousand dollars. The OIR had to employ more and more student research assistants in order to assist with the administration of the surveys as well as the resulting data entry.

Discovering the potential of a web-based system

Several factors came into play as the OIR determined that the development of a system to administer surveys on the web was a priority. First, as mentioned above, the office had used two different scanning mechanisms with the hope that data could be entered faster and reduce the need for manual data entry. The first scanning system ("system" meaning the use of specific software with a stand-alone scanner) was implemented in 1996, with the goal to scan many of the larger surveys while only making minor adjustments to the formats of the actual instruments. The procedure of reformatting the instruments turned out to be quite labor intensive for the research analysts, and the actual scanning turned out to be problematic, with much time spent on cleaning up the resulting data files. In 1998 a new scanning software was purchased along with a higher-grade scanner in order to correct the problems encountered with the previous scanning system. But, while the software was a bit more robust and the scanner faster, there was still an inordinate amount of time being spent on the back-end cleaning up data files. While the OIR's Executive Director and other staff members realized that a more cost effective way of administering surveys was important, it was determined that the old method of manual data entry was more efficient, and by the summer of 1999 there was virtually no scanning of surveys in the OIR.

Also, the staff of the OIR was becoming aware of the viability of conducting web-based surveys through sources within the field of institutional research and through new technology being utilized externally, especially in the business sector. By 1999 the use of web-based surveys in institutional research was being discussed widely and actually being used by a few institutions. There seemed to be little consensus, however, on how effective it was (or would be) in addressing the data collection needs of an institutional research office. But, since the OIR at Tufts had already been thinking strategically about changing processes – as culminating in the

scanning efforts – the idea of doing surveys via the web was taken seriously by the executive director and staff.

The technology was improving quickly, and by this time commerce and data collection through the web was standard operating procedure by businesses. The advancement of Java and other graphical interfaces made it easier to format and collect data, and the improvements in security gave people a greater sense of confidentiality. But, most importantly, it made interaction user-friendly for the respondent.

Selection of web-based software

The vision the office had in selecting the software and hardware for web-based surveying focused on maximizing survey effectiveness and making the entire survey administration process efficient. To do this, the ultimate goals were to: 1) implement a system that was seamless, 2) to have control over and ensure the security of the process and resulting data, and 3) to have a methodologically sound process.

Seamless. Generally, all quantitative data analysis in the OIR is conducted using one specific statistical analysis package. One concern that was persistent among staff members was the form data would take when it is compiled from a web-based survey. While there are generally standard methods to convert data from flat files or other forms to the data analysis file, there were concerns that the process would be time consuming, adding to the research analysts' already burgeoning workload, and that it would increase the likelihood of corrupted data. Keeping these concerns in mind, a decision was made early in the process to test the new web survey software package put out by the same company that produces the data analysis software. This way, seamlessness would be paramount: while the survey is being developed the data file, with variable and value labels, is also being developed; and, the data that is collected from survey submissions would be automatically placed into the data file, alleviating concerns about conversion.

Control. Another concern staff members had was the ability to control the process as well as the data. Actually, there was concern at two different levels. First, it was deemed important that the surveys stay on a campus-owned server, since there is sometimes proprietary data or sensitive data derived from the surveys conducted by the OIR. There was a fear that if the survey did not reside on a university server the owners of a private server – which are commonly used for different web survey packages – could collect ID's or email addresses of respondents and use them for their own purposes or sell them. There was also concern over sensitive data, especially admissions-related, falling into the hands of competing institutions.

Second, staff members thought it would be important that members of the OIR have direct control over the placement of the surveys on a server and extraction of the data once it is submitted. This was due to the fact that analysts are often under tight time frames within various stages of the survey process, and relying on a server administrator for such custodial actions may conflict with the need to place and replace surveys on the server in an instant.

Methodologically sound. Of course, the most important aspect of selecting a software package was its ability to allow the OIR to create and administer surveys that collect reliable data. One important element was to have the ability to create survey items in several different formats, based on the nature of the inquiry. The ability to combine good methodology with seamlessness and control in a software package would be the challenge of selecting the right product.

Selection

Early discussions with campus information technology professionals focused on the choice between developing our own program or purchasing a package. Creating a program had been done on other campuses using HTML for design and PERL or Cold Fusion for security and data collection, but there was no expertise in the OIR to complete such a task. After describing exactly what was needed, it was the consensus of the IT people that: 1) it would take a great deal of time and effort to build such a program, 2) they did not have the resources to support this type of initiative, and 3) that we should explore "off the shelf products."

Based on the OIR's needs and expectations of a web-based software package, it was decided to test the web survey package put out by the same company that produces the office's data analysis software. After consulting with the company's sales representative concerning its attributes and viewing a live demonstration, it was determined that this product would make the process efficient and seamless. In keeping with the objectives of the project, it would allow the office to place it on a university server (but not control its custodianship since the OIR did not have its own server); and it was methodologically sound. The product was then forwarded to the university's Information Technology Services (ITS) unit for testing. The web-based software package consisted of two primary components: the actual survey development software and its own server software, which needed to reside on a platform server with certain size attributes. It was ITS's responsibility to test both the development software and server component for compatibility with the university server. This testing was extremely important since this was the first version of the product. ITS informed us that it was compatible with their servers and that they could support its use if we decided to purchase it. So we did.

Implementation

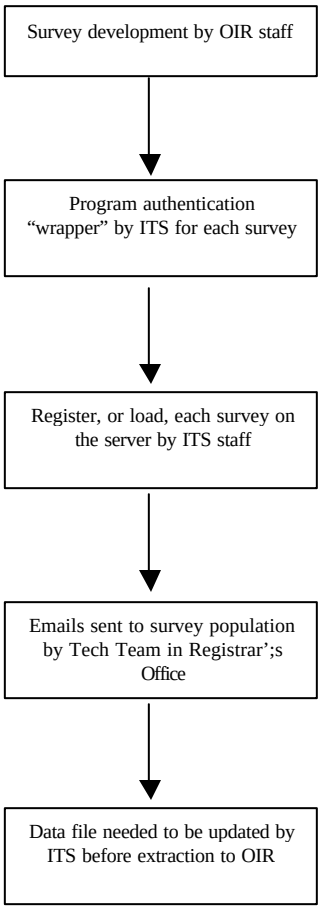
First Product

It was decided that the first survey to be administered on the web by the OIR would be the 2000 Pre-Major Advising Survey, which is distributed to Sophomores in the Arts and Sciences and First-Year students in Engineering. These students are required to complete the survey in the spring semester in order to be able to register for fall courses. Thus, it was important to ensure that only those students who were eligible to complete the survey did, and that their survey completion was verified in order to release their registration hold. The server component was placed on the university server in the Fall of 1999 and OIR staff began developing the actual web-based survey with the survey development component.

The server component, however, did not have an authentication function, so the first unexpected task was to build one to function with this product. Since the product selected was Java-based, an authentication function could not be built right into the server, and an electronic "wrapper" had to be built that could authenticate a user before accessing the survey. Since the OIR did not have access to the university server, staff members administering web surveys now had to rely on ITS professionals to build authentication programs for each survey (when needed) and load the survey on the server. The process was now getting more complicated and the critical path of survey administration was being extended, as shown in Figure 1.

With an authentication program now written, the next step was to load and register the survey on the server. This also became an extended and iterative process. In order to receive data into a data file, a master file had to be registered on the university's

Figure 1. Critical path of web-based survey administration: Initial phase



network server. There were several technical problems, however, relating the registration of the master file. Some of the problems were on the server end, and some were on the survey development end. By January of 2000 after several weeks of struggling with the first version of the product, a second version was released and the OIR was selected as a beta site. The

second version was to have addressed many of the server interaction problems ITS was experiencing. While expectations were high that the process of loading and registering surveys would now be smoother, initial results were not positive. Compounding the difficulties was the departure of the ITS professional who had been assigned to the project.

While the new ITS contact was a very good programmer, he needed time to come up to speed on the project. After experiencing more difficulties with registering the master file and loading the survey, a successful load and registration was finally completed. In March of 2000 a test survey was on the server and receiving data. Since sophomores and first-year engineers started registering for fall classes in late March, there was very little time to do extensive testing on the survey. Staff members and student research assistants in the OIR and ITS staff members submitted test data. Test data was collected and no serious problems reading the form or submitting responses were identified. Campus-wide testing was not done.

When the survey went live and students were directed to complete the survey, many were unable to access and complete it. While there were some problems with identification numbers in the authentication process, the main problem identified was that students were unable to load the survey on their browsers or that it froze or crashed while they were completing it. The OIR then had to supply paper versions of the survey to those who could not load the survey. Of the 1,400 eligible students to complete the survey, a total of 866 submitted responses via the web, for a web response rate of 62%. Another 101 students completed paper surveys, for a combined response rate of 69%. Usually around 10% do not register for the fall of their Junior year due to study abroad plans, so there was still about 20% who did not complete the survey who still needed to register for classes. A determination was made by members of the registrar's staff to allow students to register regardless of whether or not they had completed a survey. Thus, due to the difficulties in reading and loading the survey the OIR lost up to 27% of their potential web responses.

Due to the loss of so many respondents, other web-based projects in the pipeline were put on hold until an evaluation of the pre-major advising experience could be conducted and procedures/modifications found to improve the process. When it was determined that a detailed training session on the software package – for both OIR staff and ITS staff – would be most appropriate, a two day session in the Fall of 2000 was scheduled with technical staff members of the company that produces the product. The goal of the training was to cover the survey process in a comprehensive manner, including survey development and server interaction.

After the training was completed the OIR decided to proceed slowly in administrating web surveys. The Residential Life Survey, which is administered in January of each year to all students living in campus residence facilities, would be the next survey to go on-line, followed by an evaluative survey of the new student services center. The process of loading and registering the survey went smoother than the Pre-Major Advising survey. There was also less feedback concerning problems with loading and submitting the survey by respondents. But, even though incentives were offered for completing the surveys, there was no requirement to complete these surveys. Thus, as the OIR staff learned later, many students did have trouble loading the

survey, but since it was not required, like the Pre-Major Advising survey, many simply disregarded it and did not make the effort to contact the OIR with problems. The final response rate for Residential Life was 34%, about 15% lower than previous paper administrations. The response rate for the student service center survey was, at 17% , lower than expected.

To learn more about the lower response rates, a focus group was held with undergraduate students who had recently received emails announcing the residential life and student services center surveys. Most had not completed either. While they prefer to complete web surveys as opposed to paper surveys, they generally are not motivated to complete any survey unless it is compelling. They felt that a well-defined subject heading on the email announcing the survey is important, but are not necessarily motivated by prizes unless it is monetary in value. One issue that did not come out of the focus group was technical difficulty in loading and submitting the surveys

In the meantime, however, as the Residential Life Survey was being administered and plans to do the Pre-Major Advising Survey (renamed Sophomore Survey) were being discussed, the OIR staff began to explore alternative packages for conducting web surveys. While progress was being made in utilizing the current software package, the OIR staff came to a consensus that there were three problems in using the product: 1) the survey development part of the program was somewhat labor intensive and clumsy; 2) the need to have ITS create wrappers and register each survey was not very efficient; and, most importantly, 3) there seemed to be no clear answer to resolving the problem of survey compatibility with a diversity of browsers on the respondents' end. Multiple contacts with the company's technical support did not provide an answer. The OIR became keenly aware of this problem when the Sophomore Survey was administered in March. When students called the OIR to report problems, they were asked what browser they were using and if they had enabled Java in the browser. Table 1 highlights the browser versions that generally worked and those that did not[1].

Table 1. Browser version compatibility with initial web surveying program

Survey worked	Survey did not work
Netscape 4.5	Netscape pre-4.5
Netscape 4.7	Netscape 6.0
Exlorer 5 installed after October, 1999	Explorer 5 installed before October, 1999
	Any Mac

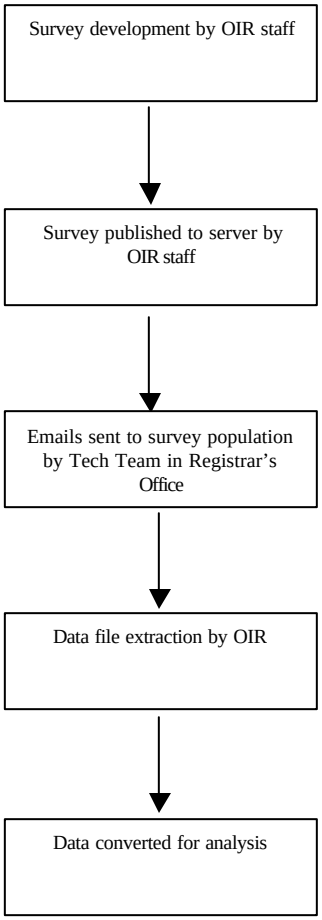
Second product

1 These were general trends based on information supplied by the respondents.

With these three problems in mind, the OIR staff began collecting information from both campus-based IT consultants and professionals in the institutional research field with experience or knowledge of web surveying to find out how to address these problems. An analysis by a consultant at Tufts concluded that the only way to comprehensively address these problems was for the OIR to have its own server, which was not feasible due to the need for additional staff to run and maintain it..

Many alternative products were suggested by other IR professionals, and several were tested. The product that was selected for a trial run addressed all three of the necessary conditions cited by the OIR staff. First, it was very easy to use in the development stage. The question editor is very intuitive and display templates save considerable time in formatting. Second, the surveys reside on a private, non-university server but the user has complete control over the publishing and republishing of surveys to the server. No intermediary technical staff person is needed. The company also ensures, in writing, the confidentiality and security of the data. And third, while this product is also Java-based, there are no browser compatibility problems. Unfortunately, one goal the OIR did not reach in implementing this second product was seamlessness. While there are extra steps needed to convert the raw data to data that will reside in the standard data analysis software, those extra steps are well worth the effort based on the time and effort saved on the front end of the process. The critical path now involves less players and places more control in the hands of the OIR, as shown if Figure 2.

Figure 2. Critical path of web-based survey administration: Second phase



Conclusions

The process of selecting hardware and software to maximize survey effectiveness as well as to make the entire survey administration process efficient was a complex task. The institutional research office had three goals: 1) to implement a system that was as seamless as possible – that is, to integrate survey development, distribution, collection and analysis into an efficient process that minimized the critical path of survey administration, 2) to have institutional control over the process and resulting data, and 3) to implement a survey design and process that was methodologically sound. Analysis of the entire selection and implementation process shows that, based on the technical support and technical infrastructure available, seamlessness and control eventually had to be compromised in order to maintain methodological credibility.

Implications

The implications of this case study are great for those institutional researchers looking to migrate from paper surveys to web-based surveys. The lessons learned from this case study are numerous. It is critical to fully understand the level of technical support and complexity in your office and on your campus. In order to successfully implement a web-surveying system, the product chosen must meet your support needs. Technical support is critical. Make sure the technical support that is provided by the vendor is timely. Just because you do not hear about the problems it does not mean that they do not exist. Pilot testing was created for a reason. Invest time in exploring options. Be patient, go slow. Once you figure out how to best implement web-surveys on your campus, the results are astounding.

References

Underwood, D., Kim, H. & Matier, M. (2000). To mail or to web: Comparisons of survey response rates and respondent characteristics. Paper presented at the Annual Forum of the Association of Institutional Research, Cincinnati, OH, May 21-24, 2000.

Cartwright, D., Thompson, R., Poole, M. & Kester, D. (1999). Assessing distance learning using a website survey. Paper presented at the Annual Forum of the Association of Institutional Research, Seattle, WA, May 30-June 3, 1999.

Matz, C.M. (1999). Administration of web versus paper surveys: Mode effects and response rates (Master's thesis, University of North Carolina-Chapel Hill, 1999). EDRS, ED 439694.

Porter, S & Umbach, P. (2000). Surveys of students and faculty: Using good practices and the Internet to lower costs and increase response rates. Workshop held at the annual conference of the Northeast Association for Institutional Research, Pittsburgh, PA, November 2000.

Smith, C. (1997). Casting the net: Surveying an Internet population. Journal of Computer Mediated Communications, 3(1) 77-84.

COMPARISON OF STUDENT SUCCESS AMONG ASNUNTUCK COMMUNITY COLLEGE ELEMENTARY ALGEBRA STUDENTS PLACED BY ACCUPLACER SCORE, SCHOLASTIC APTITUDE TEST SCORE, OR PREREQUISITE COURSE

Patricia Hirschy, Associate Professor of Mathematics
Qing Mack, Director of Institutional Research
Asnuntuck Community College

Introduction

Asnuntuck Community College is one of 12 colleges in the Connecticut Community College system. It is located in north central Connecticut, 25 miles north of Hartford, Connecticut, and 5 miles south of Springfield, Massachusetts. Asnuntuck has a service area of eight towns. In Fall 1999, 1,538 students were enrolled in more than 200 credit courses at Asnuntuck. Ninety percent of the students were from the eight service towns, and 3% were from the neighboring state of Massachusetts. Asnuntuck is primarily a liberal arts college, and it offers 18 two-year degrees and 18 certificates. The college employs 24 full-time faculty and approximately 80 adjunct faculty. Sixty-one percent of the students are female, 24% are full time, 76% are part-time, and 68% are 22 years or older.

MATH 101 (Algebra I) is an elementary algebra course offered by Asnuntuck for students who are not yet prepared to take a college level mathematics course. Some of these students have never had an algebra course, other students were not successful in previous algebra courses, and still others took algebra so long ago that they have forgotten their skills and need a review of the concepts. The purpose of this course is to offer students the opportunity to study or to review elementary algebra skills that are a foundation for college mathematics courses such as contemporary mathematics, elementary statistics, and intermediate algebra.

Students are placed into Algebra I through one of three policies. One policy is based on the Accuplacer test score. The Accuplacer test is a computerized placement test with subtests in mathematics, English, and reading. Students who earn a high enough score on the arithmetic and algebra subtests can be placed directly into Algebra I. A second placement method is based on the Scholastic Aptitude Test (SAT). Students who score 400 or higher on the mathematics portion of the SAT can be placed directly into Algebra I without first taking the Accuplacer test. The third placement method is based on success in the prerequisite mathematics course. The prerequisite course for Algebra I is Prealgebra, a course with arithmetic and beginning algebra content. Students who have successfully completed Prealgebra or its equivalent with a grade of C or higher are eligible to enroll in Algebra I.

Purpose of the Study

The purpose of the study is to determine whether students placed by three different

methods-Accuplacer score, SAT score, or prerequisite course-have different success rates in Algebra I. Comparing the success rates of students will provide an indication of the effectiveness of the placement methods, since they are established to identify those students whose mathematics background is appropriate for Algebra I. If the results of the study indicate that the student's mathematical skills are overestimated, then Algebra I will be at too high a level and the student will struggle to be successful. For this student, a better placement recommendation would be the prerequisite course. If the results of the study indicate that the student's mathematical skills are underestimated, then Algebra I will be at too low a level. The student will not be challenged in the course, and the resulting lack of motivation may affect the student's behavior and performance. In this case, a better placement recommendation for this student would be the next level mathematics course.

Review of the Literature

Students arriving at the community college door are often unprepared for college level classes (McCabe, 2000, p. 4). Forty-one percent need remediation in at least one of the basic disciplines: reading, writing, and mathematics. Twenty-five percent of the entering community college students have deficiencies in reading, and twenty percent have deficiencies in writing. Mathematics is the area with the greatest deficiency, with 34 percent of community college students needing remediation.

There are a number of reasons why students arrive at college without the necessary skills to do college-level work (McCabe, 2000, p. 39). In many states, high school exit competencies do not match entrance competencies for college. Also, high school students may be counseled into outdated general or occupational curricula. Further, mature students who begin college many years after high school graduation need courses to refresh their academic skills. A larger percent of high school graduates are continuing their education with post-secondary studies (Davis, 1989, p. 22). The increased demand for higher education resulted in a proliferation of open admissions colleges. The government has increased its financial support of developmental education, thus making it a more attractive offering.

Along with the increased number of unprepared students came a need for placement policies to determine whether students had the skills to be successful in college-level courses. According to the American Mathematical Association of Two-Year Colleges (2001), the purpose of placement is to place students in accord with their educational goals and prerequisite knowledge (p. 1). The college policies should be applied equally to all student placements, and multiple measures should be included in the process. Colleges should validate the effectiveness of the placement policies and should evaluate them on an ongoing basis.

Wattenbarger and McCleod (1989) investigated a variety of mathematics placement measures that have been used by community colleges. Statistical tests of the relationship between student course grades and standardized college examinations such as the SAT and the ACT resulted in low correlations (p. 18). The SAT and ACT are intended as measures of general academic aptitude, so using them for placement decisions may be inappropriate (p. 20).

Content-oriented tests would provide information that is more useful for placement decisions. Also, combining other factors such as high school mathematics background, student motivation, and student learning style would result in a more effective placement policy.

Research indicates that mandatory testing and placement is an essential component of effective education (McCabe, 2000, pp. 42-45). It is important that colleges have information regarding the nature and extent of student academic needs so the appropriate support services can be developed and offered. Students enrolled in classes without the requisite skills are in danger of failing those courses. Instructors face students with widely differing skill levels, thus making it difficult to balance the needs of unprepared students with the needs of the course.

In a review of ten exemplary college developmental education programs (McCabe & Day, 1998, p. 22), comprehensive placement and assessment efforts were critical factors of success in every program. Bucks County Community College, for example, has established mandatory placement and assessment policies to fulfill its goal to ensure that students "are provided with an opportunity to begin their studies at an appropriate level" (p. 38). Prince George's Community College administers standardized testing to guarantee that students have a foundation for college-level instruction (pp. 79-83). As part of the process, the college offers a placement test confirmation procedure that includes a second departmental skills assessment and an intensive mathematics review course.

Procedures

Data Collection

The study sample comprised Asnuntuck students who enrolled in all sections, day and evening, of Algebra I in Fall 1999 and who persisted sufficiently long to receive a course grade. Students who were "No Show" or who withdrew within the first two weeks of class were not included in the sample. An ex post facto research design was used, since the data was based on the records of students who have already enrolled in Algebra I and have received a grade. The information was retrieved from the Banner Student Record System, from current student folders, and from archived student folders.

Demographic and academic information was collected on each student including gender, age, major, method of placement, Accuplacer score, SAT score, Prealgebra course grade, Algebra I course grade, number of credits completed prior to Fall 1999, and enrollment status (full-time or part-time).

Data Analysis

This research study applied both descriptive and inferential statistics. Descriptive statistics were used to organize and present the demographic information such as age, number of credits completed, and gender. Inferential statistics were used to determine if there is a relationship between success in the course and placement methods. A Chi-square test was applied to determine if there was a statistically significant difference in success rates among students placed by Accuplacer score, prerequisite course, or by SAT score. The course grade

was chosen as the measure of success since there are no standardized tests or departmental finals that would yield a different measure of student success.

Results and discussion

Of the 148 students who registered for the course, n = 107 students were eligible for the study. Forty-one students were not eligible due either to missing information or to placement other than by the three defined methods. A comparison of the study sample to the general college population showed that students who took Algebra I tended to be younger, enrolled full-time, in their first year of college, and more likely to be matriculated in a degree or certificate program (Table 1). The average age for all Asnuntuck students in Fall 1999 was 32 years old, compared to 25 for those who took Algebra I. Fifty-three percent of the Algebra I students were fulltime, compared to 24% for the general college population. Eighty-nine percent of the Algebra I students were first-year students with 12 or fewer credits, compared to 75% of the general college population. Fifty-nine percent of the Algebra I students were matriculated, compared to 45% of the general college population.

TABLE 1
Demographic/ Academic Information
Of
Asnuntuck Students and Sample Population

	Total	Full-Time	Part-Time	Male	Female	Under 22	22 and Older
All College	1538	371	1167	598	940	490	1048
% of Total	100%	24%	76%	39%	61%	32%	68%
M101 Students	107	57	50	41	66	60	47
% of M101	100%	53%	47%	38%	62%	56%	44%

	Total	1st time	Matriculated	1st Year	2nd Year
All College	1538	498	694	1151	387
% of Total	100%	32%	45%	75%	25%
M101 Students	107	44	63	95	12
% of M101	100%	41%	59%	89%	11%

Students placed by the SAT score (n = 20) tended to be traditional age and enrolled in day sections. Students placed by the prerequisite course tended to be non-traditional age and enrolled in evening sections. Of the twelve second-year students, ten students were placed by the prerequisite course and two by Accuplacer score.

The Chi-square test was first applied to the 107 students in the study. The independent variable was placement method-SAT score, prerequisite course, or Accuplacer score. The dependent variable was success in course, where success is defined by C grade or higher. As can be seen from Table 2, there is no statistically significant difference in the data. Of note, though, is the range of success rates, from a low of 43.2% from prerequisite course placement

to a high of 60% from SAT score placement.

TABLE 2

M101 Attendees by all Three Placement Methods

			M101 grade by group		
			C or better	C- or lower	Total
M101 Attendees	W/SAT	Count	12	8	20
		% within M101 Attendees	60.0%	40.0%	100.0%
	Took M100	Count	16	21	37
		% within M101 Attendees	43.2%	56.8%	100.0%
	P/M101	Count	26	24	50
		% within M101 Attendees	52.0%	48.0%	100.0%
Total		Count	54	53	107
		% within M101 Attendees	50.5%	49.5%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1.546[a]	2	.462
Likelihood Ratio	1.554	2	.460
Linear-by-Linear Association	.084	1	.772
N of Valid Cases	107		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 9.91.

The Chi-square test was then applied to student groups categorized by age, gender, course section (day or evening), and student status (first or second year). In some cases, the Chi-square test could not be performed due to low cell values. This difficulty arose for evening students placed by SAT score, for female students placed by SAT score, for traditional students placed by prerequisite course, for non-traditional students placed by SAT score, and for second-year students. For the remaining tests (listed below), there was no statistically significant difference at the 0.05 level.

- All Algebra I sections
- All day sections
- Evening students placed by Accuplacer and prerequisite course
- Female students placed by Accuplacer score and prerequisite course
- Traditional age students placed by Accuplacer and SAT score
- Non-traditional students placed by Accuplacer and prerequisite course
- First-year students

Even though there was not a statistical significance, there were two comparisons of note. For traditional students in day classes, those placed by SAT score had a much higher

success rate of 62.5% compared to those placed by Accuplacer score with a 32% success rate. This resulted in a P-value of 0.055-higher than the 0.05 level of significance but close enough to warrant further investigation (Table 3).

TABLE 3

Traditional Age Students Enrolled in Day Classes

			M101 grade by group		
			C or better	C- or lower	Total
M101 Attendees	W/SAT	Count	10	6	16
		% within M101 Attendees	62.5%	37.5%	100.0%
	P/M101	Count	8	17	25
		% within M101 Attendees	32.0%	68.0%	100.0%
Total		Count	18	23	41
		% within M101 Attendees	43.9%	56.1%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3.685[b]	1	.055		
Continuity Correction[a]	2.551	1	.110		
Likelihood Ratio	3.713	1	.054		
Fisher's Exact Test				.105	.055
Linear-by-Linear Association	3.595	1	.058		
N of Valid Cases	41				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.02.

Also, for students who were placed by prerequisite course (37 students), female students achieved a 51.9% success rate compared to the 20% success rate earned by male students. This result may be influenced by the fact that only ten male students were placed by the prerequisite course and only two passed Algebra I with a C grade or higher.

Students Placed by Prerequisite Course

			M101 grade by group		
Sex			C or better	C- or lower	Total
Female	Took M100	Count	14	13	27
		% within useable sample by plc	51.9%	48.1%	100.0%
Male	Took M100	Count	2	8	10
		% within useable sample by plc	20.0%	80.0%	100.0%

The Chi-square test results of no statistical significance indicate that the three placement methods may be equally effective for the groups of students described in the Results and Discussion section. This shows that their chances of earning a C grade or higher in Algebra is not related to the method of placement. However, there are two groups of students for whom the analysis indicates a need for further investigation. Traditional students in day classes should be investigated to determine why students placed by SAT score were more successful than students who were placed by Accuplacer score (Table 3). Similarly, students placed by prerequisite course should be studied to determine why female students were more successful than male students (Table 4). In addition, this study was not able to confirm the effectiveness of placement methods for student groups with low sample sizes.

Several recommendations can be made as a result of this study. Twenty-eight students who enrolled in the Fall 1999 Algebra I course were placed by alternative methods. These placements should be studied to identify patterns that led to student success and might be considered as additional placement procedures. The demographic and academic attributes of the students placed by SAT score should be studied to identify factors that could be considered in addition to the SAT score to better predict success in Algebra I. The curriculum of both Algebra I and Prealgebra should be reviewed to determine if there are curriculum changes that could increase the success rate for students who are placed by prerequisite course into Algebra I. Gender differences for this group, in particular, should be studied.

Next Steps

Analysis of the Fall 1999 Algebra I data for all 148 students should be continued to identify factors other than placement that might be used to predict success in Algebra I. Similarly, the data analysis should continue to identify factors that might flag students who are at-risk in Algebra I. Next, data should be collected that tracks the Fall 1999 Algebra I students over the Spring 2000, Summer 2000, Fall 2000, Spring 2001, and Fall 2001 terms to determine if further mathematics courses were taken, when the students took the courses, and how well the students performed compared to other students in these courses. Persistence rates and enrollment patterns (continuing enrollment, stop out, transfer, drop out, and graduation) should be included as part of the investigation.

References

American Mathematical Association of Two-Year Colleges. (2001, September). Draft position statement of the American Mathematical Association of Two-Year Colleges on Initial placement of community college students into the mathematics curriculum. Memphis, TN: Author.

Davis, R. M. (1989). A Curriculum in flux: Mathematics at two-year colleges. Washington, D. C.: Mathematical Association of America.

Isaac, S., & Michael, W. B. (1997). Handbook in research and evaluation (3rd ed.). San Diego, CA: Educational and Industrial Testing Services.

McCabe, R. H. (2000). No one to waste. Washington, D. C.: Community College Press.

McCabe, R. H., & Day, P. R. (Eds.). (1998). Developmental education: A twenty-first century social and economic imperative. Mission Viejo, CA: League for Innovation.

Wattenbarger, J. L. & McLeod, N. (1989, Spring). Placement in the mathematics curriculum: What are the keys? Community College Review, 16 (4), 17-21.

ASSESSING THE ECONOMIC IMPACT OF A UNIVERSITY ON THE STATE AND COMMUNITY

Heather Kelly Isaacs
Institutional Research Analyst
University of Delaware

Abstract

Economic impact studies are utilized to demonstrate the beneficial effects that colleges and universities have on the local economy. In fall 1999, the University of Delaware conducted such a study to determine the economic impact that student, faculty and staff, and University expenditures have on the local community and state. This paper discusses the methodology and results of this study in detail.

Introduction

Economic impact studies are utilized by higher education institutions to demonstrate financial importance to their respective local communities. Stout suggests that colleges and universities exchange costs for benefits with their communities and economic impact studies help institutions to maximize their perceived financial benefits (1998). During fall 1999, the Office of Institutional Research and Planning at the University of Delaware (UD) conducted an Economic Impact Study. The purpose of this study was to determine the economic impact that student, faculty and staff, and University expenditures have on the local community and state. A survey was administered to the students and faculty and staff on the Newark campus, as well as local businesses in the Newark community. The final research report for this study contains three main sections that discuss the economic impact of University students, faculty and staff, and local businesses on Newark and the state of Delaware.

Developing the Study

A number of institutions have previously conducted economic impact studies (i.e., James Madison University, Pennsylvania State University, Southeastern Louisiana University, Universities of Arizona, Florida and Massachusetts-Amherst, Utah Valley State College, and Virginia Commonwealth University). These studies were reviewed before initiating the study at the University of Delaware. The study at UD borrowed the survey instruments utilized by Southeastern Louisiana University (Baldwin, Boeckman, & McKenzie, 1998). Modifications were made to these survey instruments to incorporate additional variables. A pilot test of the revised survey instruments was conducted among students and faculty and staff before the surveys were distributed to the actual sample population. In addition, conversations involved professors of economics at the University to ensure that the methodology and survey instruments were sound. One professor indicated a concern that while the survey instruments

were comprehensive, people might feel that they require too much effort to complete. However, individuals who participated in the pilot test did not comment on the lengthiness of the survey instruments. Comments and suggestions from the pilot test and various conversations were incorporated and the final survey instruments were developed.

While the survey instruments would provide primary data on the direct expenditures of students and faculty and staff, it was also necessary to obtain secondary data for the University. The secondary data provided information on the direct expenditures and revenues of the University. The secondary data were obtained by contacting University departments such as conference services, dining services, athletics, special events, and purchasing.

The method for this particular economic impact study is based on the Caffrey-Isaacs model developed for the American Council of Education (ACE) in 1971. This method, also known as the ACE method, is the most widely used to determine the economic impact of higher education institutions (Stokes & Coomes, 1998). It generates a series of impact indicators on the basis of simple linear cash-flow formulas. The first step of the ACE method is to identify the expenditures of the college community, including direct institution spending to vendors and local spending of students, faculty and staff, and visitors. The student expenditures should not include direct payments to the institution for tuition, housing, and food. A regional economic multiplier is then applied to the total impact of local spending by the institution, students, faculty and staff, and visitors. Another component of the ACE method is that the labor market impact is estimated by the increase of jobs in the area. It should be noted that Caffrey and Isaacs do not distinguish between the expenditure impacts of resident and non-resident students. The economic impact study developed at the University of Delaware utilizes many of the concepts of the ACE method.

Methodology

During the months of October and November 1999, the Economic Impact Study questionnaires were administered to students and faculty and staff on the Newark campus, as well as to local businesses. A follow-up mailing was conducted in February. The student questionnaire was administered to approximately 2,800 undergraduate and graduate students at the University. The students were randomly chosen to ensure a representative sample by class level, ethnicity, gender, campus status, residence status, and time status. The original student data set contained 688 surveys. The student response rate was approximately 25%. The final student data set was weighted during the analysis process to correctly represent the overall percentages of students by class level, gender, campus status, residence status, and time status. The weighted data thus provide findings from the sample of students to represent the Newark campus undergraduate population.

The faculty and staff questionnaire was administered to approximately 1,820 faculty and staff members on the Newark campus. The faculty and staff members were randomly chosen to

ensure a representative sample by both employment and time status. The original faculty and staff data set contained 938 surveys. The faculty and staff response rate was approximately 52%. The final faculty data set was weighted during the analysis process to correctly represent the overall percentages of faculty and staff members by employment and time status. The weighted data thus provide findings from the sample of faculty and staff to represent the Newark campus faculty and staff population.

The business questionnaire was administered to the business owners and managers of approximately 270 local businesses in the Newark area. The businesses contacted were located on Main Street and approximately a five-mile radius to the University. The business data set contained 90 surveys. The business response rate was approximately 34%.

Findings

Student Economic Impact

The mean monthly student income from all sources after taxes was approximately $1,020. The total mean monthly student expenditures spent in Delaware were approximately $780. Student expenditures ranged from housing to entertainment to medical and dental. Please note that students were asked to exclude University tuition, housing, and meal plans from their monthly expenditures.

Student expenditures spent in Delaware varied by gender, time status, class level, residence status, and campus status. The total mean monthly expenditures spent in Delaware were approximately $860 for female students and $690 for male students. Female students spend a larger percentage of total monthly expenditures than their male peers on housing, utilities, telephone and cable, clothing, other retail, and medical and dental. Male students tend to spend more on entertainment and recreation, services, and automobiles. The percentage of total monthly expenditures for all other categories was equal.

The total mean monthly expenditures spent in Delaware were approximately $520 for full-time students and $1,880 for part-time students. This large expenditure difference can be attributed to the fact that part-time students tend to be older and are most likely employed full-time. The percentage of total monthly expenditures that full-time students spend on telephone and cable, food and beverage, entertainment and recreation, clothing, and books and educational supplies is greater than their part-time peers. Part-time students tend to spend more on housing, utilities, services, other retail, automobiles, and medical and dental.

The total mean monthly expenditures spent in Delaware by class level increased linearly. The largest expenditure category for all class levels except freshmen and sophomore students was housing. The housing expenditure for freshmen was extremely low due to the fact that 89% of freshmen live in University-approved housing. Freshmen tend to spend the greatest

percentage of their total monthly expenditures on food and beverage (20%) followed by books and educational supplies (17%). Sophomores tend spend the greatest percentage of their total monthly expenditures on automobiles (20%) followed by food and beverage (18%) and housing (17%). After housing expenditures, seniors, juniors, graduate students, and continuing education students tend to spend the greatest percentage of their total monthly expenditures on food and beverage and automobiles.

The total mean monthly expenditures spent in Delaware were approximately $1,040 for resident students and $490 for non-resident students. This large expenditure difference may be attributed to the fact that the total mean monthly expenditures are greatest for continuing education students and 79% of these students are Delaware residents. The percentage of total monthly expenditures that non-resident students spend on housing, telephone and cable, food and beverage, entertainment and recreation, clothing, and books and educational supplies was greater than their resident peers. Resident students tend to spend more on utilities, services, other retail, automobiles, and medical and dental.

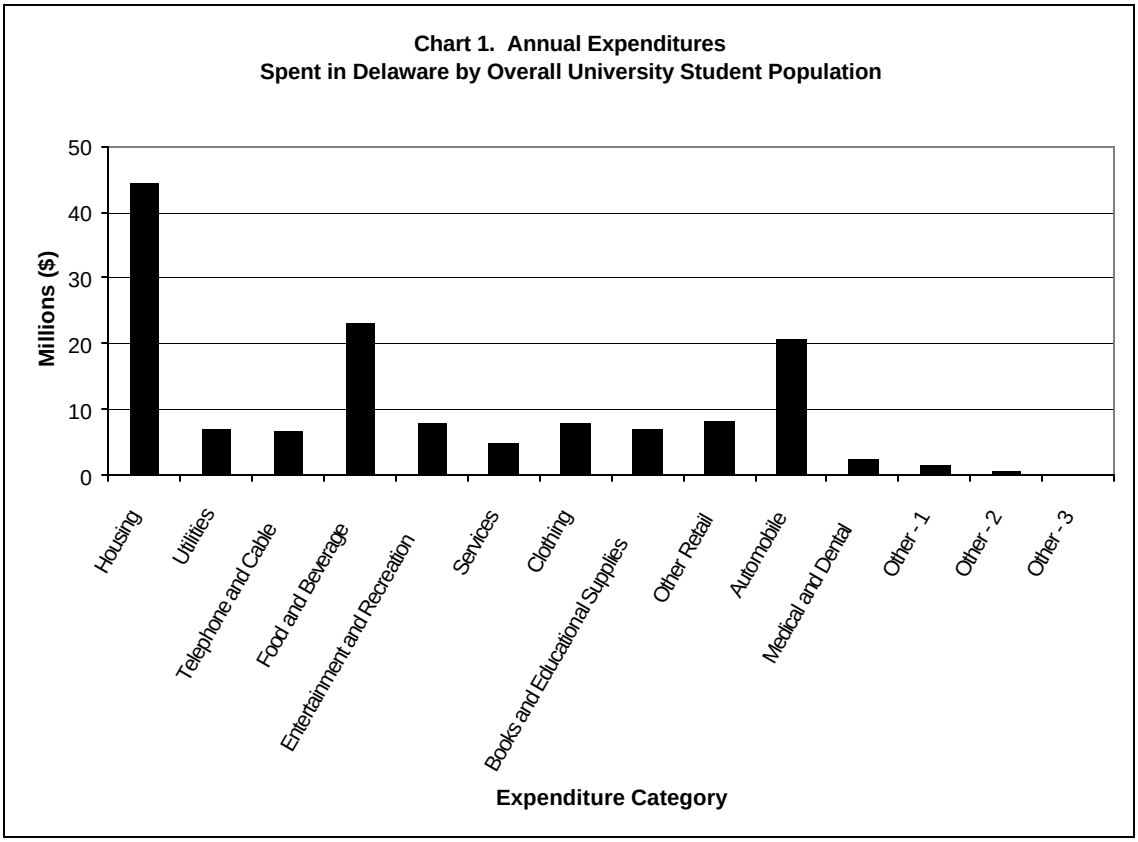
The total mean monthly expenditures spent in Delaware for on-campus students were approximately $250 and $1,100 for off-campus students. This large expenditure difference may be attributed to the fact that the on-campus students have minimal housing and utility expenditures. The percentage of total monthly expenditures that on-campus students spend on telephone and cable, food and beverage, entertainment and recreation, services, clothing, books and educational supplies, and other retail was greater than their off-campus peers. Off-campus students tend to spend more on housing, utilities, automobiles, and medical and dental.

The estimated total annual expenditures spent in Delaware by the overall University student population were approximately $143,003,950[1] (see table 1 and chart 1). The breakdown of these annual expenditures is summarized on the following page.

[1] The annual expenditures for each category were calculated by multiplying the monthly expenditure by the student headcount for each term by the number of months in each term. The terms (number of months) included fall 1999 (4), winter (1), spring 2000 (4), and summer 1 and 2 (1.5 each). The total annual expenditures were the sum of these categories.

Table 1
Annual Expenditures Spent in Delaware by Overall University Student Population

	Expenditures Per Year ($)[1]	% of Total Annual Expenditures
Housing	44,506,332	31.1
Utilities	7,113,717	5.0
Telephone and Cable	6,748,911	4.7
Food and Beverage	22,982,778	16.1
Entertainment and Recreation	7,843,329	5.5
Services	4,924,881	3.4
Clothing	8,025,732	5.6
Books and Educational Supplies	6,931,314	4.8
Other Retail	8,208,135	5.7
Automobile	20,793,942	14.5
Medical and Dental	2,553,642	1.8
Other – 1	1,641,627	1.1
Other - 2	547,209	0.4
Other - 3	182,403	0.1
Total Annual Expenditures	***143,003,952***	***100***



In addition, students were asked to indicate up to six (6) Newark businesses that they frequent on a regular basis. Of the top 23 businesses that students mentioned, more than half (57%) were in the food and beverage industry, 30% were retail stores, 9% were grocery stores, and 4% provided general entertainment (i.e., video store).

Faculty and Staff Economic Impact

The total mean monthly faculty and staff household expenditures spent in Delaware were approximately $2,320. Faculty and staff expenditures ranged from housing to retail to education and tuition. Faculty and staff household expenditures spent in Delaware varied by employment status, residence status, and state of residence. Faculty members tend to spend the most in the state of Delaware, followed by professional staff, hourly staff, and salaried staff. The largest expenditure category for all employment status groups except hourly staff is housing. After housing expenditures, faculty, professional and salaried staff tend to spend the greatest percentage of their total monthly expenditures on food and beverage followed by automobile. The largest expenditure category for hourly staff is automobile. After automobile expenditures, hourly staff tend to spend the greatest percentage of their total monthly expenditures on housing followed by food and beverage.

The total mean monthly expenditures spent in Delaware for resident faculty and staff members were approximately $2,690 and $1,120 for Delaware non-resident faculty and staff members. Delaware residents tend to spend the greatest percentage of their total monthly expenditures spent in Delaware on housing (32%) followed by food and beverage (16%) and automobile (14%). Delaware non-residents tend to spend the greatest percentage of their total monthly expenditures spent in Delaware on food and beverage (21%) followed by education and tuition (17%) and automobile (12%).

The total mean monthly faculty and staff expenditures spent in Delaware varied by state of residence. Delaware residents spend the most in the state of Delaware, followed by Maryland, Pennsylvania, and New Jersey residents. Delaware residents tend to spend the greatest percentage of their total monthly expenditures spent in Delaware on housing (32%) followed by food and beverage (16%) and automobile (14%). Maryland residents tend to spend the greatest percentage of their total monthly expenditures spent in Delaware on education and tuition (20%) followed by food and beverage (19%). Pennsylvania residents tend to spend the greatest percentage of their total monthly expenditures spent in Delaware on food and beverage (27%) followed by automobile (16%) and education and tuition (13%). New Jersey residents tend to spend the greatest percentage of their total monthly expenditures spent in Delaware on other retail (30%) followed by automobile (29%) and food and beverage (23%).

The estimated total annual expenditures spent in Delaware by the University Newark campus faculty and staff population were approximately $94,501,840[2] (see table 2 and chart

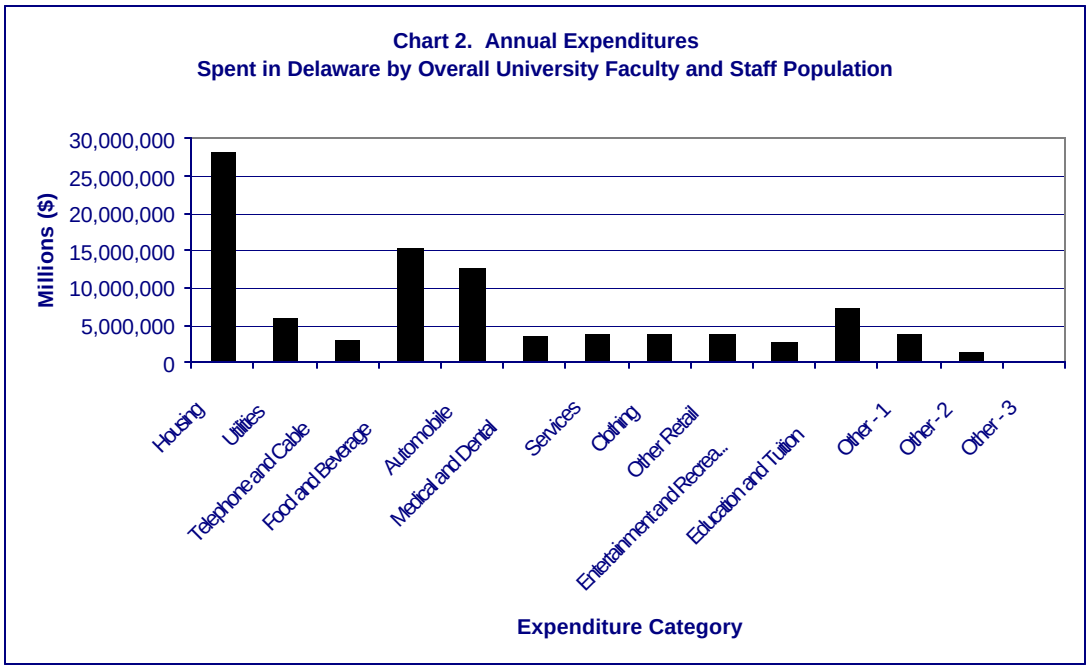
[2] The faculty and staff annual expenditures were based on the Newark campus only (n=3,393).

2). The breakdown of these annual expenditures is summarized below:

Table 2

Annual Expenditures Spent in Delaware by Overall University Faculty and Staff Population

	Expenditures Per Year ($)[2]	% of Total Annual Expenditures
Housing	28,094,040	29.7
Utilities	6,066,684	6.4
Telephone and Cable	3,012,984	3.2
Food and Beverage	15,227,784	16.1
Automobile	12,540,528	13.3
Medical and Dental	3,420,144	3.6
Services	3,745,872	4.0
Clothing	3,705,156	3.9
Other Retail	3,664,440	3.9
Entertainment and Recreation	2,687,256	2.8
Education and Tuition	7,247,448	7.7
Other - 1	3,745,872	4.0
Other - 2	1,302,912	1.4
Other - 3	40,716	0.04
Total Annual Expenditures	***94,501,836***	***100***



University Economic Impact on Local Businesses

Local businesses indicated that they employ a number of current University students and alumni. They also indicated that they benefit from revenues generated by University students and faculty and staff. A number of the local business respondents indicated that the University and its community influenced decisions regarding products and services. In addition, the business respondents reported that the University and its community influenced decisions regarding scheduling events and sales. Other business decisions that are influenced by the University and its community included advertising, hiring, and business location choice. A number of the business respondents indicated that the University and its community were an asset to their business. In addition, the respondents indicated that the University had a positive influence on the Newark community and business sales. One respondent indicated that the University brought "vibrant" life to the community. Respondents also indicated that the University of Delaware made Newark "work" and the effect of the University and its community on their business as a whole was overall positive. Business respondents indicated that the faculty and staff at the University were a positive influence on their enterprise and that the University was important in making their operation successful, a driving force in the market share, provided competition, and encouraged a diverse population.

Overall University Economic Impact

University Revenues

The University's largest source of operating revenue was tuition and fees. In addition to operating revenue, the University generated revenue through special events and activities. For example, during the 1998-99 fiscal year, Clayton Hall hosted a number of meetings and events both internally and externally. Approximately 58% of the events hosted were external. The external meetings and events included the following types: corporate, associations, government, non-profit, public relations, religious, social, and education. In addition, during the 1998-99 fiscal year, 246,221 visitors attended the Bob Carpenter Center for intercollegiate athletic events, trade shows, and concerts. Eleven of the events at the Bob Carpenter Center consisted of two comedy shows, two children's shows, one family show, one specialty show, and five concerts. Approximately 48,965 visitors attended these 11 events and approximately $1,127,800 was generated in revenue.

University Expenditures

The University of Delaware is the 8th largest employer in the state of Delaware. During fall 1999, the University employed approximately 3,400 faculty and staff members on the Newark campus. The University compensated these employees approximately $154,775,980.

The University makes a number of purchases through both Delaware and non-Delaware vendors. During the 1998-99 fiscal year, the University purchased approximately $62,835,400 worth of products and services through Delaware vendors. Purchasing in Delaware accounts for 41% of the University's overall purchasing.

Economic Impact Summary

The direct expenditures of students, faculty and staff, and the University account for a large part of the University's economic impact on the state of Delaware. These direct expenditures lead to indirect purchases where additional services are purchased, employees are paid, and these employees, in turn, make additional expenditures. This is called the "multiplier" effect. The "multiplier effect" has also been defined as the ratio of increased income to increased spending (Stokes & Coomes, 1998). The total economic impact of the University of Delaware was calculated by applying a multiplier of 1.9[3] to the direct expenditures. The estimated total economic impact of student and faculty and staff direct expenditures and University purchasing is summarized below:

Table 3

Annual Expenditures Spent in Delaware by the University of Delaware and Its Community

	Estimated Spending in Delaware Per Year	Overall Economic Impact
Student Expenditures	$143,003,952	$271,707,509
Faculty and Staff Expenditures	$ 94,501,836	$179,553,488
University Purchases	$ 62,835,388	$119,387,237
Total Economic Impact	$300,341,176	$570,648,234

During 1999, the University and its community spent approximately $300 million in Delaware. These estimated expenditures spent in Delaware ($300 million) are 3 times the state operating appropriations level ($90 million). The estimated total economic impact of the University of Delaware is approximately $570 million.

The University of Delaware is also responsible for generating additional jobs for businesses that provide goods and services to the University and its community. The Bureau of Economic Analysis suggests that approximately 36 jobs are generated for each additional $1 million dollars of output[3]. Looking at the estimated student and faculty and staff expenditures and University purchases in the state of Delaware, approximately 10,810 new jobs are generated which

3 U.S. Department of Commerce – Bureau of Economic Analysis. (1992). *Regional multipliers: A user handbook for the regional input-output modeling system (RIMS II)*. 3rd ed. Washington, D.C.: U.S. Government Printing Office.

increases the overall economic impact of the University of Delaware.

Conclusion

This study provides evidence that the University of Delaware's impact on the local community and state through economic benefits is immense. The students and the faculty and staff contribute a great deal to the local and state economy through their personal and household expenditures. Local businesses indicate that the University and its community are a positive influence and contribute to their success. Overall, the return on the state's investment in the University of Delaware is approximately 3 times greater than its initial investment.

In addition to economic benefits, the University provides a number of additional benefits to the local community and the state of Delaware. These include, but are not limited to, employment opportunities, social and cultural events, educational opportunities, and community development.

Conducting an economic impact study has provided the University with the necessary means to effectively communicate the value of the University in economic and social terms. This proves to be very useful when interacting with decision makers that include, but are not limited to, government officials, local businesses, and the local community.

Acknowledgements

I would like to take this opportunity to thank and acknowledge Donald Boeckman of Southeastern Louisiana University who granted permission to the University of Delaware to adopt their methodology and survey instruments in order to develop this Economic Impact Study.

REFERENCES

Applied Technology Research Corporation. (1998, April). *Southeastern Louisiana University: Community impacts.* Baton Rouge, LA: Author.

Baldwin B., Boeckman, D., & McKenzie, L. (1998, May). *Estimating a university's economic and community impact: Principles, procedures, and outcomes*. Paper presented at the meeting of the Association for Institutional Research, Minneapolis, MN.

Stokes, K., & Coomes, P. (1998). The local economic impact of higher education: An overview of methods and practice. *AIR Professional File*, 67, 1-14.

Stout, R. (1998, May). *The effectiveness of economic impact studies by liberal arts colleges: A case study*. Paper presented at the meeting of the Association for Institutional Research, Minneapolis, MN.

INSTITUTIONAL RESEARCH AND EFFECTIVENESS: WHERE RESEARCH MEETS STRATEGIC PLANNING

Ana Lucia Kazan-Fishman, Ph.D.

Director, Institutional Research & Effectiveness
Lehigh Carbon Community College, Schnecksville, Pennsylvania

Introduction

The changing role of the institutional researcher: From statistics expert to people expert

"Institutional researchers have long struggled with the definition of their profession, or indeed, whether what they do may be considered a profession" (Huntington & Claggett, 1992). From the office of statistics institutional researchers are moving toward more dynamic roles such as those related to planning and administrative decision-making (Banks and Colby, 1989).

Notwithstanding the differences in scope, staff size and qualifications between institutional research offices in two and four-year colleges, we all still have to address reporting requirements and accountability issues pertaining to the competitiveness of the high education market place (Cyphers, 2001). More and more institutional researchers are actively participating in strategic planning processes and implementation, becoming involved in college-wide values assessment, mission, vision, goals, and action strategies determination.

Purpose of this paper

Within the framework described above, this paper intends to present the case of the Lehigh Carbon Community College, which has just gone through its fifth strategic planning effort, the first ever conducted by the director of institutional research and effectiveness. The experience described in this paper, the responsibility to conduct or at least actively participate in their institution's strategic planning process, is becoming more and more common for institutional researchers. In this paper the author presents a five-step strategic plan she adapted and the process she developed for implementing it on a larger scale, i.e., that of a community college.

Developing a Strategic Plan Tailored to a Community College

The Lehigh Carbon Community College

Lehigh Carbon Community College (LCCC) was born Lehigh County Community College in 1966, sponsored by the Lehigh and Carbon counties school districts. In 1994, the name was officially changed to Lehigh Carbon Community College. Today LCCC has graduated almost 16,000 students, approximately 30% of the total enrollment during its 35

years of existence. LCCC has an average of 4,000 credit students per semester, and an average of 14,000 credit-free students per academic year, making its annual population an estimated crowd of 22,000 students.

LCCC's Office of Institutional Research has been in existence since 1988, initially to provide statistical information to internal and external constituencies, mainly its sponsoring High School Districts, the Pennsylvania Department of Education, and the federal Integrated Postsecondary Education Data System. Always in compliance with the Middle States Association of Colleges and Secondary Schools requirements for accreditation, LCCC has taken every step necessary to improve its effectiveness process. One of these steps is the involvement of the Office of Institutional Research in the institutional planning process.

Planning and strategic planning efforts at LCCC

Prior to 1983, strategic planning at LCCC was limited. The dean of instruction, who formed an agenda for curricular matters and grant applications, and the dean of business affairs, who focused on physical plant issues, did most of institutional planning. In December 1983 the president created a planning group called Horizons Council, consisting of faculty, administrators, classified staff and trustees. The Horizons purpose was to provide leadership and direction to the college's planning effort through participation from all three of the personnel groups in addition to the trustees. Students joined the council in 1984. The Horizons Council developed a first five-year planning document, updated every six months and finalized annually. Progress toward the objectives was monitored throughout and recorded in a year-end report. A major problem with this format was that there was no formal structure for the college community to participate in the review of objectives developed by the Horizons Council. Key segments of the college, such as the Business office, academic areas, and the Office of Institutional research were underrepresented. Distribution of the progress reports generated by the Horizons Council was limited to the president's council, deans, and directors.

In 1991, the Board of Trustees decided that LCCC needed a more comprehensive strategic planning process that included input from external constituencies. A Strategic Planning Committee was formed and initiated the first major strategic plan undertaking at LCCC. Through in-depth personal interviews with representatives of all LCCC's sectors, as well as leaders of local industries, businesses, professions, agencies and organizations, the college published in 1992 its *LCCC 2000: A Strategic Vision for Lehigh County Community College*. Despite the effort, many faculty and staff members felt that too much of that strategic plan content came from a small group of people and the overall feeling of ownership was limited. In 1993 LCCC's new president replaced the Horizons Council with the college-wide Strategic Planning Committee, with the intention of revising the plan every two years through community interviews.

In 1994, the coordination of the Strategic Planning Committee fell under the jurisdicition of the recently created position of Vice-President for Research, Planning, and Community-Based Education. As an effect of changes and re-structuring, the Strategic Planning Committee

met only a few times over a period of two years, and the first update of *LCCC 2000* was not published until 1995: *LCCC 2000 and Beyond*. Although the document contained updated background information and some changes in format, the basic findings and strategic directions remained unchanged. Numerous efforts continued to be made by LCCC to conduct an effective and inclusive strategic planning process.

In 1995 planning sessions included representatives of LCCC internal and external communities; and in 1996 and 1997 planning forums involving teams of faculty, staff, students and trustees were held to formulate priorities for the next two-year period. Overall LCCC planning process was structured in the following manner: members of the Board of Trustees participated in a weekend retreat at the beginning of every academic year to discuss plans and overall strategies and goals for the institution. The president then developed his personal goals and objectives based on *LCCC 2000 and Beyond*. The president's goals and objectives were then distributed to the staff, and subsequently each department and each staff member would formulate its goals and objectives aligned with the president's and college's determined path.

This formal process, however, did not promote cohesiveness within the college. The majority of faculty and staff felt confused about the college's planning efforts since so many changes had occurred in the years before. Faculty in particular were unsure about their role in planning and whether their views were really being considered. Trying to resolve the confusion, in 1996 college administrators reorganized the standing committee structure, revamping some committees and adding new ones. One of the new committees created then was the Planning and Budgeting Committee, with representatives from faculty, administration, classified staff, and students, whose mission was to review and to make recommendations about matter related to the planning and budgeting processes.

In 1998 LCCC started a new strategic planning effort, with the newly formed Strategic Planning Steering Committee. With representatives of all areas of the college, and led by an external consultant, the effort resulted in the establishment of the *1998-2003 Strategic Plan*. The *1998 Plan* covered five main areas of the institution: (1) programs and services, (2) systems and processes, (3) people, (4) finances, and (5) facilities. This plan endured until 1999-2000, when the president of the college initiated his retirement process. After a national search the new president, Mr. Donald W. Snyder, was hired in 2000. During the period from 1995 to 2000 the college experienced a sharp decrease in enrollment and retention, and one of the new president's first measure was to revamp and upgrade LCCC efforts in recruiting and retaining students, as well as promoting internal cohesiveness by implementing a new extensive and inclusive strategic planning effort. For this effort, LCCC counted on its recently hired Director of Institutional Research, whose job title is now Director of Institutional Research and Effectiveness.

Institutional Research and Effectiveness at LCCC

LCCC Office of Institutional Research was vacant during the 1999-2000 period, being

temporarily occupied by the Enrollment Director to fulfill state and federal reporting needs. In the fall of 2000 SCT-Banner was implemented college-wide as the new administrative software. As the new Director of Institutional Research and Effectiveness, the author came to the college in September of 2000, just over a month after the new system had been implemented. Still starting its SCT-Banner learning curve, the whole college staff was struggling to gain competence in the new system, and yet state and federal level reports needed completion. In the midst of learning the new job and the new system, the author still had pending the task of leading LCCC's Institutional Effectiveness efforts, which would lead to the college's compliance with the requirements of an upcoming Middle States Accreditation review.

The model used at LCCC was based upon and adapted from the works of Bryson (1989), Lofquist (1990), Safrit (1990, 1991, 1992, and 1994), and Barry (1994). The five-steps model includes:

1. S.W.O.T. profile - when the teams determine their strengths, weaknesses, opportunities and threats
2. Determining organizational values - team members declare values that they embrace as individuals and, therefore, bring them to the team
3. Establishing team's mission - team members define their team's mission by assessing how the team contributes to the whole organization, the team's main reason for existing, the audiences the team serves, and the benefits resulting to the organization by virtue of the team members' work
4. Declaring the team's vision - the team establishes a vision, the highest standards the team wants to achieve as a group serving the large institution, by brainstorming where members think the team should go to serve the institution's future needs,
5. Determining goals, action strategies, and key performance indicators - team members decide practically and specifically what must be done to accomplish their mission and to reach their vision, as well as to determine how they get there and how they know that they have arrived.

Implementing these five steps, however, took more than only five steps. During a period of six months a foundation for the work ahead was developed: (1) a teamwork handbook; (2) the institutional effectiveness committee defined and gathered; (3) networks of information and exchange with the college's leadership team; (4) definition of teams in the midst of organizational re-structuring; (5) trainers trained; (6) activity days selected; and (7) infra-structure to support the project set in place. The process is not over yet, and to this date we still do not know how it is going to end. And that is the wonder of strategic planning.

The Process

Forming a taskforce

It is very important to grant college-wide representation in the strategic planning taskforce, i.e.,

faculty, staff, classified staff, academic and institutional administration need to actively participate in this effort. At LCCC we called the strategic planning group the *Institutional Effectiveness Task Force (IE Taskforce)*, and it included three faculty members, four administrative staff representatives, and three academic administration representatives.

Selling the idea to the taskforce

The college's Leadership Team has appointed Taskforce participants. The Leadership Team is composed of the College's president, vice-president, chief financial officer, and the deans. By virtue of being appointed and not consulted before, some IE Taskforce members were not necessarily happy to be there. Thus, the first responsibility of the Taskforce chair was to sell the idea to the other members of the taskforce. In the first LCCC IE Taskforce meeting, the Taskforce Chair brought materials to motivate and engage members in the topic, such as:

- Articles and handouts on institutional effectiveness;
- The *Teamwork Handbook*, a strategic planning model adapted for use by LCCC (see below); and
- A summary of institutional effectiveness concepts as described by our accreditation agency, the Middle States Commission on Higher Education.

The Taskforce Chair made herself available for questioning, listening to the expression of doubts and criticism, and granted everybody a good adaptation period. The LCCC Institutional Effectiveness Taskforce met every three weeks for four months. During this time, the fiercest opposition to the process came from the faculty members. They mainly expressed concerns that this process would not be effective, since is had been tried before without really changing the college. Some of this criticism focused on the process being "just another one," "an exercise in futility," "we have done this before," or "it's not going to work." It did not matter how many times the taskforce met and discussed the need for the process and the rationale behind it. Until the process had been completed, the faculty members of the taskforce were by far the toughest critics.

Gaining the support of the College's Leadership Team

Even though the process was initiated by the decision of the College's Leadership Team, it was still not clear to any of the parties, including the author's, the extent to which we needed to work collaboratively, and especially how much the Leadership Team needed to support the initiative. When it became clear to the taskforce that it would be necessary to close the college in order to conduct the many teams self-assessment and examination, we needed the Leadership Team to schedule the events. In addition, we needed the authority of the Leadership Team to meet the college-wide resistance to the process.

Defining the teams

By virtue of the administrative re-structuring, the college was going through with a new president and renewed efforts to improve enrollment and retention, administrative and academic

departments were in transition. Even the college's Leadership Team did not know yet what the final administrative format would be. After unsuccessful tentative efforts to divide the whole college into teams, and after waiting for more definitions from the Leadership Team regarding the new administrative format, the IE Taskforce finally reached agreement. Faculty and staff were divided into academic areas and in budget areas, respectively. With the help and authority of the Leadership Team, the IE Taskforce conducted separate meetings with academic area heads and account directors, when the process was explained, and copies of the Teamwork Handbook were distributed. On those occasions the IE Taskforce Chair detailed the procedures, answered questions, and made herself available for addressing specific and personal concerns that might emerge. Academic heads and account directors then became responsible for distributing amongst their team members copies of the Teamwork Handbook, as well as explaining how it should be implemented before the *D-Days* scheduled for the whole college to conduct its strategic planning. We adopted the expression *D-Days* to refer to the days in which the College would be closed for the strategic planning to be conducted.

It would be virtually impossible to close the whole college for one full day, so the IE Taskforce along with the college's Leadership Team decided to conduct the faculty and staff strategic planning processes on two different days. With summer approaching, it was becoming increasingly difficult to find a day in which the entire faculty would be able to participate. The Leadership Team decided that the afternoon of the Spring Graduation would be ideal for the Faculty *D-Day*, since traditionally all faculty participate in the graduation ceremony that would take place early that evening. The *D-Day* for staff was planned for a Friday, few weeks into the summer semester, when the movement of students was not so intense.

Faculty was divided into 11 teams:

1. Healthcare professions
2. Science
3. Mathematics
4. Human Services
5. Technology
6. Computer Science
7. Business
8. Counselors
9. Humanities
10. Social Sciences
11. Learning Assistance

Staff was divided into 19 teams:

1. Academic Administration
2. Academic Administrative Support

3. Accounting Team
4. Administrative Services (Administration, Human Resources, Safety and Security, Switchboard)
5. Continuing Education Department
6. Duplicating/ Mail/ Word Processing
7. Educational Support: Learning Assistance Services
8. Educational Support: Literacy and Job Training
9. Information and Technology
10. Institutional Advancement Team
11. Institutional Effectiveness Team
12. Learning Resource Center
13. Marketing and Community Relations
14. Operations and Maintenance
15. Sites
16. Student Accounts
17. Student Life
18. Student Services
19. Workforce Training

Training the trainers

The LCCC faculty *D-Day* was scheduled for May, and the staff *D-Day* was scheduled for June 8th. Prior to those two major days, the Leadership Team and the Provost Team (composed by the former College's provost and the deans) went through the strategic planning process. The objective was twofold: first to test the process, and second to "create" trainers who could help facilitate the May and June sessions. Also, the Institutional Effectiveness Task Force went through the process of addressing last concerns and generating trainers who understood the process and could help facilitate the sessions with faculty and staff. A tentative train-the-trainer manual was initiated but abandoned, as it proved relatively useless when compared to actually going through the process and clarifying questions on the spot.

D-days and Infrastructure

Closing the college two half-days took planning and collaboration among the Institutional Effectiveness Task Force, the Leadership Team, and the Human Resources Department. For both days, staff and faculty members were offered lunch by the College, coffee-break, and faculty members were served dinner right before the graduation, as many of them were staying at the college for the ceremony, with no time to go home between the two events. Signs were posted on the doors one week in advance to let students know that the college would be closed during specific periods. A memo, signed by the president, stressed the importance of participating and requested that those who would be unable to come should communicate in advance with the IE

Taskforce Chair. Many staff, but only a few faculty members documented previous commitments; these parties were contacted to arrange subsequent meetings for update.

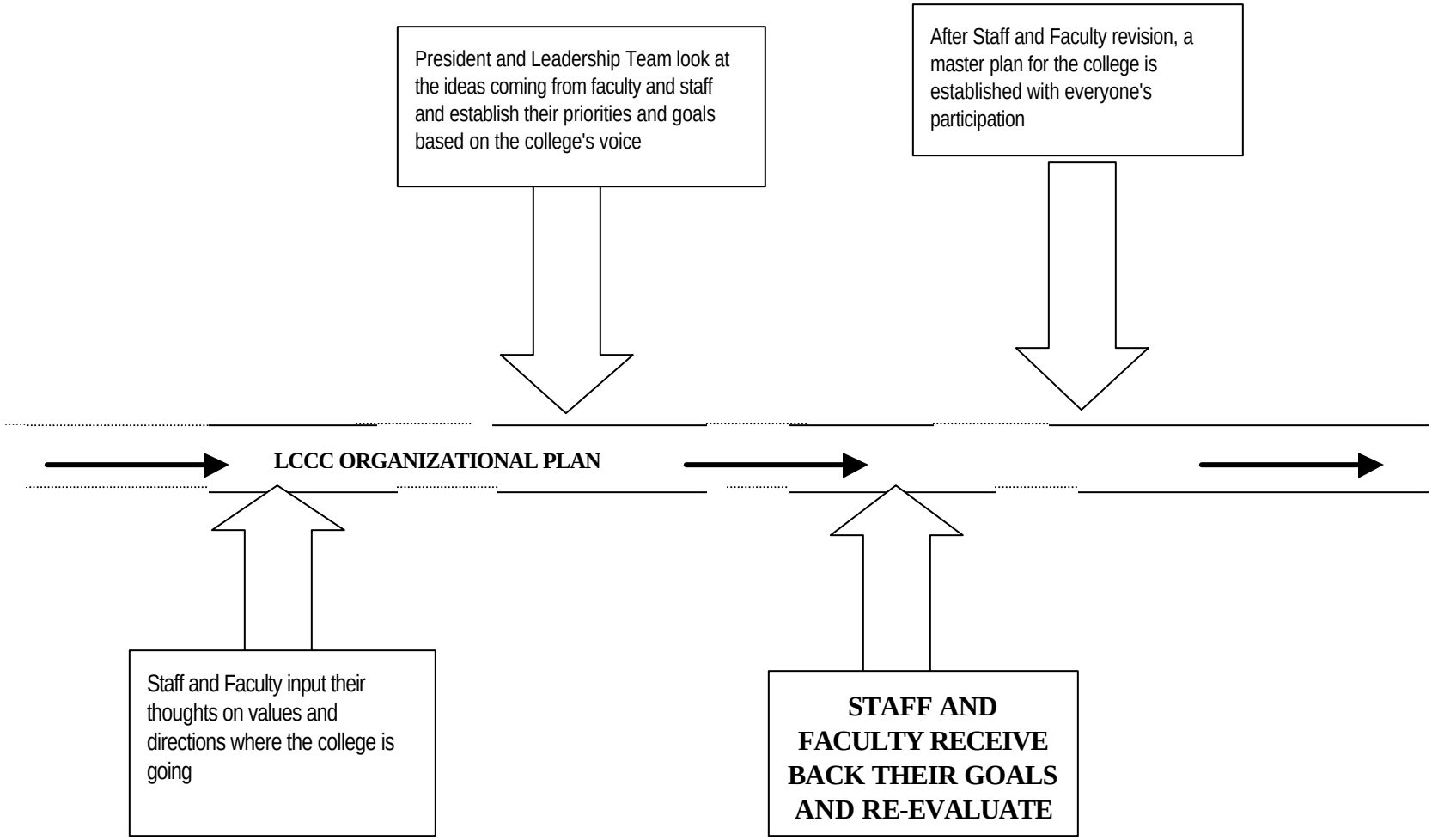
Each faculty and staff member received a memo with the agenda for their respective days. This memo also contained the names of other teammates (many of these teams were still new due to the college's re-structuring), the classroom to which they were assigned, and a reminder to bring their Teamwork Handbooks completed up to step four. The IE Task Force would have only four hours to conduct S.W.O.T. profiles, values, mission, and vision statements. Due to the limited time and impossibility of allocating more time to each day's activity, it was decided with the Leadership Team that the goals and action strategies would be completed individually by each team that could count on facilitators from the IE Task Force.

Classes were selected close to one another to facilitate the event coordination; lists of names of team members were affixed outside the door of each assigned classroom. Classrooms were equipped with easels, easel pads, colored markers, and tape so the sheets could be taped into the walls as the work progressed. A few thoughts for the day were affixed in the walls for motivation and inspiration, as suggested by a consultant psychologist. The thoughts included references to the confidentiality of the discussions that would occur during the brainstorming; to the need for overcoming a complaining mode; and to obtain help for personal difficulties with the transition period the college was going through. These thoughts were titled *Rules for the Day* and were posted on signs on the classrooms walls. Despite the good intention, these signs generated some concern among a few faculty members because of being called "Rules." As a consequence of this reaction, the IE Taskforce changed the title to *Thoughts for the Day* for the staff *D-Day*. Thus no similar concerns arose during staff *D-Day*. For both events, facilitators gathered 30 minutes before the start of activities to re-group and clarify last minute questions and concerns. No further incidents occurred, and the IE Taskforce and the college's Leadership Team considered both events successful.

Defining Goals and Action Strategies

As decided, LCCC's strategic planning process would follow the format bottom-up: top-down: bottom-up: and top-down again. The last piece of the first bottom-up part was still missing: the definition of goals and action strategies for each team. This piece would be important for the president and the Leadership Team's analysis of the college's direction for the next two-year budget period. The deadline for this analysis to start was the beginning of the Fall semester, therefore it was determined that each team should meet in their own time to go through the last step of this first phase of the process. The IE Taskforce and some account directors who have already been through the process were made available to facilitate the staff teams' processes. Since the majority of LCCC faculty is off during the Summer, it was determined that during Convocation Day, before the official start of Fall classes, academic teams would have a three-hour period to determined their goals. Once more IE Task Force members and other staff members already trained in the process would be available that day to facilitate.

LCCC STRATEGIC PLANNING MODEL



Compilation

As a result of LCCC's college-wide strategic planning process, several compiling documents were produced, with more to be produced in the future:

1. A summary of overall strengths, weaknesses, opportunities, threats, values, missions and visions of all teams for the president's analysis
2. A compilation of each teams work for the president's analysis
3. A compilation of trends of strengths and weaknesses for the president's analysis
4. A compilation of trends of opportunities and threats for the president's analysis
5. A summary to be distributed to the whole faculty and staff body of each team's mission, vision, and goals after the final input from the president and the Leadership Team (in process).

Lessons Learned

1. There is no "by-the-book" way to conduct a college-wide strategic planning in the midst of concurrent organizational change
2. Plan, plan, plan, plan, plan, plan, plan, plan ahead *(if you can*)
3. Include everybody in the planning as much as you can
4. Be very prepared for resistance – RESISTANCE IS NATURAL
5. Be prepared to act as a counselor, psychologist, confidant and pacifier
6. Assure and assure and assure everybody that everything will be all right in the end
7. Each organization has its characteristics and personality
8. Try to have advice from someone with some experience or who did it before to double check steps
9. Be prepared to improvise and adapt to specific and unexpected circumstances
10. Accept that the process will fall out of your hands at some moments and come back to you later, probably corrupted and changed
11. Be persistent, be patient

Positive Aspects of Model

1. I knew how to do it
2. I worked in many other circumstances
3. Promotes college-wide participation
4. In the back-and-forth movement, it gives opportunities for revisions

Negative Aspects of Model

1. May take too long in a larger organization

2. Gives opportunity for a lot of criticism, whining, and errors
3. May get out of hand due to the need to use several facilitators

REFERENCES

Banks, D.L. and Colby, A. (1989). Organizing institutional research in the community college. Eric Digest: Eric Clearinghouse for Junior Colleges in Los Angeles CA. Identifier: ED320648.

Barry, B.W. (1994). *Strategic planning workbook for nonprofit organizations.* St. Paul, MN: Amherst H. Wilder Foundation.

Bryson, J.M. (1989). *Strategic planning for public and nonprofit organizations*. San Francisco: Jossey-Bass Publishers.

Borg, W.R., & Gall, M.D. (1983). *Educational research: An introduction*. New York: Longman Inc.

Conklin, N.L., Jones, J.M., & Safrit, R.D. (1992, May). *Organizational values of Ohio Cooperative Extension Service employees*. Paper presented at the Symposium Research in Agriculture and Extension Education, Columbus, OH.

Conklin, N.L., Jones, J.M., & Safrit, R.D. (1992). Values: Acting on our shared beliefs. *Journal of Home Economics, 4* (Winter), 17-21.

Cyphers, C.J. (2001). *Where once was nothing: On building an office of institutional research*. Paper presented at the Forty First Forum of the Association for Institutional Research, June 3, Long Beach, CA.

Hit, W.D. (1988). *The leader-manager: Guidelines for action*. Columbus, Ohio: Batelle Press.

Huntington R.B. and Clagett, C.A. (1992). *Increasing institutional research effectiveness and productivity: Findings from a national survey*. Unpublished paper.

Jimmerson, R.N. (1989). What values will guide Extension's future? *Journal of Extension, XXVII*, 16-18.

Lehigh Carbon Community College (1997). A comprehensive self-study. Prepared for the Commission on Higher Education, Middle States Association of Colleges and Schools. Schnecksville, PA.

Lofquist, W.A. (1990). *The youth opportunity planning process: A systematic approach to involving community teams in strategic planning.* Tuscon, AZ: AYD Publications.

Pattan, J.E. (1986). The strategy in strategic planning. *Training and Development Journal*

(March), 30-34.

Rokeach, M. (1973). *The nature of human values*. New York: The Free Press.

Safrit, R.D. (1990). *Values clarification in the strategic planning process of an adult education organization*. Unpublished doctoral dissertation, North Carolina State University, Raleigh.

Safrit, R.D. (1991, October). *The role of values clarification in the strategic planning process of adult education organizations*. Paper presented at the Annual Conference of the American Association for Adult and Continuing Education, Montreal, Quebec, Canada.

Safrit, R.D. (1991, October). *Organizational values clarification as a component of strategic planning: New horizons for qualitative and quantitative evaluation*. Paper presented at the Annual Meeting of the American Evaluation Association, Chicago, IL.

Safrit, R.D. (1992). *Creating our future together: The Florida 4-H strategic planning conference, Participant handbook*. (Available from author)

Safrit, R.D. (1992). *Creating our future together: The Texas strategic planning conference, Participant handbook*. (Available from author)

Safrit, R.D. (1994, April). *Evaluating today's culture and visioning for tomorrow's growth: Effectively utilizing stakeholders in planning strategically for adult education.* Paper presented at the International Conference on the Practice of Adult Education and Social Development, Jinan, People's Republic of China.

Safrit, R.D. (1994, May). *Evaluating today's culture and tomorrow's directions: Effectively utilizing stakeholders in organizational strategic planning.* Paper presented at the 14th Annual Conference of the Canadian Evaluation Society, Quebec (Quebec).

Safrit, R.D. (1994, August). *Participatory strategic planning in the College of Food, Agricultural, and Environmental Sciences, Departmental Resource Handbook*. Columbus: The Ohio State University.

Safrit, R.D. (1999). *Charting a course: Involving stakeholders in effective strategic planning. Participant Resource Handbook*. The Ohio State University: College of Food, Agricultural, and Environmental Sciences, Department of Human and Community Resources Development.

Simerly, R.G., & Associates. (1987). *Strategic planning and leadership in continuing education*. San-Francisco: Jossey-Bass Publishers.

Tichy, N.M. (1983). *Managing strategic change: Technical, political, and cultural dynamics*. New York: John Wiley and Sons.

Tobin, G.A. (Ed.). *Social planning and human service delivery in the voluntary sector*. Westport, CN: Greenwood Press.

Teamed States Department of Agriculture - Extension Service. (1991, September). *The strategic business planning process in 4-H: National strategic action planning for 4-H youth development*. Washington, D.C.

Teamed States Department of Agriculture - Extension Service. (1991, October). *4-H national strategic business plan (first revision)*. Washington, D.C.

Williams, M.S. (1992). *Strategic planning for 4-H.* (Available from Mary S. Williams, CED III, Nassau County Extension Service, P.O. Box 1550, Callahan, Florida, 32011)

Wysner, J. (1991). *Rays of hope*. Los Angeles: Authors Unlimited.

THE IMPACT OF A SERIES OF WRITING INTENSIVE COURSES ON SUCCESS ON THE WRITING PROFICIENCY REQUIREMENT

Kevin B. Murphy
Research Analyst
Office of Institutional Research and Policy Studies
University of Massachusetts Boston

Introduction

In April 1999, the Office of Institutional Research and Policy (OIRP) received a request from the Writing Proficiency Requirement (WPR) Committee for research on the connection between the curriculum and success on the WPR. The requirement consists of the successful completion of a timed essay examination, or the submission of a portfolio of work that includes several examples of papers written for courses and a new paper based on assigned readings and specific questions. It is designed to "…assist students in acquiring critical skills. Foremost among these is the ability to present ideas clearly, correctly, and persuasively in English prose" (UMB Undergraduate Catalog). The requirement must be successfully completed as a prerequisite for graduation from the College of Arts and Sciences (CAS) and from the College of Nursing (CN). It is a high stakes requirement. There is no alternative path to graduation.

One focus was to be on a group of courses that were designed to prepare students for the WPR. These were the Core or "C" courses, which were offered in a number of disciplines throughout CAS. They were overseen by the Core Curriculum Office, which is also responsible for the administration of the WPR. The requirement called for students to complete five "C" courses. Of the five required core courses, three were to be at the 100 level, and the other two at the 200 level. A core course emphasized the nature of knowledge and the methods of investigation that characterize the disciplines within its distribution area. Core courses provided instruction and practice in such intellectual skills and habits of thought as analytical writing, critical thinking, quantitative reasoning, and research techniques. In general, students were to complete their core courses before they attempted the WPR. Transfer students with 30 or more credits were not required to fulfill the requirement. This system was changed in fall 2000. The new First Year Seminar (FYS) courses replaced the old "C" courses for newly matriculating students. Students who matriculated prior to that date who were subject to the old rule are now required to complete a total of two "C" courses at either the 100 or 200 level.

Before I could examine the relationship between the curriculum and the WPR, I needed to identify all of the steps leading up to the WPR. Therefore, I conducted a process evaluation. The results of this evaluation were presented in a previous report. However, a short summary would be valuable.

The Writing Proficiency Requirement should be viewed as a process that begins before a single course is ever taken at UMB, rather than as an event. There are well-established rules for the process. It was fairly easy to identify how the process works, or, at least, how it is supposed to work. It involved attending orientation, completing an English Placement Assessment (EPA) and receiving a placement recommendation, completing the recommended courses and the five "C" Courses, and attempting the WPR at about 60 credits. If the student was successful, there were no further related requirements. If the student was unsuccessful, there were two specialized courses to prepare the students to retake the exam, and tutoring would also be made available.

In practice however, we found that large numbers of students failed to attend orientation or to get an EPA recommendation, failed to complete the appropriate courses if they had a recommendation, and especially failed to complete all of the "C" courses that were designed to prepare them for the WPR. Because all freshmen were subject to the rule, this report will focus on the behavior of students who entered UMB as freshmen.

Of the nearly 1,000 freshmen in the study, barely 17% utilized the system in the manner in which it was designed as may be seen in Table 1.

Table 1: Freshman Compliance with the WPR Preparatory Process

Group	Number of Students	Percent
All first time freshmen	993	100.0%
Freshmen with an EPA	638	64.2%
Freshmen who complied with the EPA	595	59.9%
Freshmen who complied with the EPA and completed at least 5 "C" courses	170	17.1%

Analyzing the Impact of the "C" Courses

Given the poor level of compliance with the system, the question of whether completion of "C" courses is important remained. This analysis will examine the impact of "C" courses for all freshmen, regardless of whether they had an EPA recommendation or whether they complied, if they had one. Significance tests were run on the differences between the number of "C" courses completed by students by whether they had an EPA recommendation, whether they complied with that recommendation, and whether they had one and complied with it versus those who either did not have an EPA or had one and did not comply with it. The differences were not significant in any case, and are presented in Table 2.

Table 2: Significance Test Results on "C" Courses by EPA Status

Comparison Groups	Mean	Difference	|T| Value	Probability >|T|
EPA	3.152	.008	.0683	94.55%
No EPA	3.144			
EPA and complied	3.149	.08	.2856	77.53%
EPA Did not comply	3.069			
EPA and Complied	3.149	.006	.0562	95.52%
No EPA /did not comply	3.143			

Given these small differences, it seems reasonable to examine the impact of the number of "C" courses alone without considering EPA compliance status.

The number of "C" courses taken before attempting the WPR varied considerably, with less than 30% of the students having completed at least five of the "C" courses prior to attempting the WPR. The specifics are presented in Table 3, which follows.

Table 3: "C" Courses completed by Freshmen Prior to the First WPR Attempt

'C" Courses Completed	Number of Students	Percent of Students
0	101	10.2%
1	101	10.2%
2	156	15.7%
3	158	15.9%
4	198	19.9%
5	235	23.7%
6	37	3.7%
7	7	0.7%

Methods

The data used for this study come from official University of Massachusetts Boston files. Prior to June of 1996, only data for the most recent attempt was maintained on the database. Therefore, we would were unable to tell when courses were completed in relation to the first attempt if more than one attempt was needed. The system was changed for the June 1996 examination. Therefore, the analysis group is limited to students (freshmen in this study) who attempted the WPR for the first time between June 1996 and June 2000 inclusive. Because of the attendance patterns of UMB students, this group included students who entered UMass Boston as freshmen as early as fall 1984 and as late as fall 1999.

No attempt was made to account for the quality of the students' "C" course experiences. This study is not intended to be an evaluation of the program as a whole. The analysis does not deal with assessment of program activities or of classroom implementation of those activities (Hughes). It simply focuses on the relationship between the students' performance on the WPR and completing some number of "C" courses. Any final grade that carried credit was counted as the successful completion of the course. No attempt was made to control for the discipline in which the course was offered. For the purposes of this study, all "C" courses were created equal.

Comparison of means tests and simple bivariate logistic regression models will be used. The dependent variable will be the result of pass (1) or fail (0) on the first attempt on the WPR. The independent variable will be the number of "C" courses successfully completed by the student prior to that first attempt. The overall first attempt pass rate was 76.7%. Students who completed six or seven "C" courses will be folded into a category of five or more completed "C" courses.

The analysis has been conducted using Stata for Windows® Version 6.0

Results

Even UMB students who entered as freshmen are not a homogenous group. About 58% of the students in this study were female, 36% were over age 25 (including several who were over age 60), and the group is racially and ethnically diverse with significant numbers of international students. They came to the university with varying levels of preparation. Many are non-native English speakers. Nevertheless, the first step was to examine the relationship between the "C" courses and success for the group as a whole. The observed values for all freshmen entrants are presented in the following table.

Table 4: Observed Pass Rates for All Freshmen by Number of "C" Courses

Courses Completed	0	1	2	3	4	5 or More
Pass Rate	72.3%	75.3%	72.4%	75.3%	76.8%	82.1%
N Size	101	101	156	158	198	279

No particularly strong pattern is observed except that the largest jump in pass rates is for those who completed four courses versus those who fully complied with the program by completing at least the five recommended courses. Based on this, a new dummy variable called "full_c" was created that had a value of 0 if the student completed four or fewer courses and 1 if the student completed five or more. A comparison of means test was run on the pass rate variable by the full_c variable. Those who did not complete the full sequence (N=714) passed at a rate of 76.4%, while those who did (N=279) had a pass rate of 82.1%. The difference of just over 7% returned a T-statistic with an absolute value of 2.4957, which was significant at above the 95% level (98.73%).

The more interesting question is whether the "C" courses have an incremental effect rather than being successful for only those who complete the program. Here a bivariate logistic regression was run. The results are presented in Table 5 below. The overall model returned a chi square statistic of 5.55. The probability of a larger chi square statistic by chance alone is about 1.85%.

Table 5: Logit Estimates of Pass 1st by Number of "C" Courses

Pass 1st	Coefficient	Standard Error	Z	P>|Z|	95% Confidence Interval
"C"Courses	.1007818	.0427616	2.357	0.018	.0169706 - .184593
Constant	.8846124	.1481727	5.97	0.000	.5941993 - 1.175025

Stata allows one to predict the values of the dependent variable by variation in the independent variables. In Table 6 which reports those predicted values, we can see that, on average, each "C" course successfully completed prior to the first attempt at the WPR increases the probability of passing by about 2%.

Table 6: Predicted Pass Rates for All Freshmen by Number of "C" Courses

Courses Completed	0	1	2	3	4	5 or More
Pass Rate	70.8%	72.8%	74.8%	76.6%	78.4%	80.3%

While small, this 2% increment in the pass rate for all freshmen has some importance. If all of the students had completed the full program, we would predict that about 35 more of them would have passed the test on the first attempt. This would reduce the expenses associated with a second (or subsequent) WPR administration, of operating the two specialized support courses, and of the tutoring and other administrative supports that is offered to support those who have failed the WPR.

Given the diversity of our student body, it is possible that completing "C" courses has more of an effect on some groups than on others. The first group I wanted to examine were the students who entered through the DSP program. The Directions for Student Potential (DSP) program is a free six-week, pre-admission summer program that provides academic advising, career planning, and personal counseling. DSP offers intensive workshops in reading, writing, mathematics, and study skills. Those students who successfully complete the program are admitted to the College of Arts and Sciences in the fall. It is designed for those students who show academic promise, but do not meet the traditional admission criteria.

The mean number of "C" course completed was not significantly different for DSP and non-DSP students. The observed pass rate by number of "C" courses is presented in Table 7, which follows.

Table 7: Observed Pass Rates for DSP Entrants by Number of "C" Courses

Courses Completed	0	1	2	3	4	5 or More
Pass Rate	35.7%	48.4%	59.4%	62.5%	65.1%	75.5%
N Size	14	31	32	24	43	53

In this case, the strength of the pattern is immediately noticeable. A comparison of means test was run on the pass rate variable by the full_c variable. Those who did not complete the full sequence (N=144) passed at a rate of 56.9%, while those who did (N=53), had a pass rate of 75.5%. The difference of just over 18.5% has an associated T-statistic with an absolute value of 2.3974, which was significant at above the 95% level (98.25%).

The logistic regression was then run. The results are presented in Table 8, which follows. The overall model returned a chi square statistic of 10.08. The probability of a larger chi square statistic by chance alone is about 0.15%. Only the number of "C" courses was significant.

Table 8: Logit Estimates of Pass 1st by Number of "C" Courses for DSP Students

Pass 1st	Coefficient	Standard Error	Z	P>|Z|	95% Confidence Interval
"C"Courses	.2730294	.0880809	3.100	0.002	.1003941 - .4456648
Constant	-.341596	.3000381	-1.139	0.255	-.9296598 - .2464678

Once again, the predicted pass rates were obtained, and are reported in Table 9.

Table 9: Predicted Pass Rates for DSP Entrants by Number of "C" Courses

Courses Completed	0	1	2	3	4	5 or More
Pass Rate	41.5%	48.3%	55.1%	61.7%	67.9%	74.6%

Using the predicted values of both of the previous regressions and all other things being equal, the DSP students who take no "C" courses can expect their pass rates to trail those of the overall group who took no "C" courses by about 29.3%. However, those DSP students who complete the full sequence would expect their pass rates to lag those of all students who complete the full sequence by only 5.7% and to lag the overall group average by just over 2%. These differences suggest that completing the full sequence of "C" courses was particularly important for students who entered through the DSP program.

Part of this may have been because of native language status. In this analysis, we used two separate indicators for native language status. One is "ESL". While some people object to the term "ESL", here it has a specific meaning. ESL students are students whose English language skills needed enough additional work that their EPA recommendation was for a sequence of English as a Second Language courses, or who did not have such a recommendation but took ESL courses anyway. Overall, about 10% (N=100) of all freshmen were ESL.

The other language group was the non-native English speaking students. I will follow the practice of Hamp-Lyons (1996) and use NNS for non-native speakers and NS for those who are native speakers of English. This group is very difficult for us to identify because we have no flag for it in our computer systems. Non-native English speakers were identified by an ESL recommendation or course, the presence of a score for an ESL assessment, the presence of a TOEFL score, or a recommendation for one of the English composition courses designed specifically for non-native English speakers. While all ESL students were non-native English speakers, many non-ESL students were NNS. Overall, about 19.8% (N=197) of all freshmen were NNS. We understand this to be a serious undercount. Results from the 2000 administration of the National Survey of Student Engagement and from two surveys administered in fall 2000 and spring 2001 in UMB First Year Seminar (FYS) courses indicate that the true percentage of NNS in our student population is probably closer to 40%. Separate regressions were run for the NNS students and for the ESL subset of NNS.

A comparison of means test was run on the number of "C" courses completed by NNS vs. NS students. On average, NS students completed 3.3 courses while NNS students completed only 2.7 courses. This difference of about .6 of a course had an associated T-statistic with an absolute value of 4.5955, which falls significantly above the 99% confidence level. This difference becomes important if there is a positive relationship between completing the courses and success for NNS students.

Table 10: Observed Pass Rates for NNS students by Number of "C" Courses

Courses Completed	0	1	2	3	4	5 or More
Pass Rate	50.0%	55.9%	59.1%	63.3%	66.7%	81.3%
N Size	34	34	44	49	42	48

In this case, the strength of the pattern is also immediately noticeable. A comparison of means was run on the pass rate variable by the full_c variable. Those NNS students who did not complete the full sequence (N=203) passed at a rate of 59.6%, while those who did (N=48) had a pass rate of 81.3%. The difference of over 21.5% has an associated T-statistic with an absolute value of 2.8389, which was significant at above the 99% level (99.51%).

The logistic regression was then run. The results are presented in Table 11 below. The overall model returned a chi square statistic of 9.67. The probability of a larger chi square statistic by chance alone is about 0.19%. Only the number of "C" courses was significant.

Table 11: Logit Estimates of Pass 1st by Number of "C" Courses for NNS Students

Pass 1st	Coefficient	Standard Error	Z	P>|Z|	95% Confidence Interval
"C"Courses	.2453326	.0803607	3.053	0.002	.0878284 - .4028367
Constant	-.0782742	.2433763	-0.322	0.748	-.5552829 - .3987346

Once again, the predicted pass rates were obtained, and are reported in Table 12.

Table 12: Predicted Pass Rates for NNS Students by Number of "C" Courses

Courses Completed	0	1	2	3	4	5 or More
Pass Rate	48.0%	54.2%	60.2%	65.9%	71.2%	76.3%

Using the predicted values of the previous regressions, the NNS students who take no "C" courses can expect their pass rates to trail those of the overall group who took no "C" courses by about 22.8%. However, those DSP students who complete the full sequence would expect their pass rates to lag those of all students who complete the full sequence by only 4% and to lag the overall group average of 76.7% by just 0.4%. While these differences are statistically significant, they suggest that completing the "C" courses was also particularly important for NNS.

A comparison of means test was run on the number of "C" courses completed by ESL vs. non-ESL students. On average, non-ESL students completed 3.2 courses while ESL students completed only 2.4 courses. This difference of about .8 of a course had an associated T-statistic with an absolute value of 4.8008, which falls significantly above the 99% confidence level.

Table 13: Observed Pass Rates for ESL Students by Number of "C" Courses

Courses Completed	0	1	2	3	4	5 or More
Pass Rate	40.0%	36.8%	45.5%	53.3%	64.7%	75.0%
N Size	15	19	22	15	17	12

In this case, the strength of the pattern is also immediately noticeable. A comparison of means was run on the pass rate variable by the full_c variable. Those ESL students who did not complete the full sequence (N=88) passed at a rate of 47.7%, while those who did (N=12) had a pass rate of 75%. The difference of over 27% has an associated T-statistic with an absolute value of 1.7833, which was not statistically significant (P>|T|=92.24%). Because of the very small numbers, statistical significance is hard to attain, but the numbers still bear reporting. It is troubling that only 12% of the ESL students completed the full sequence of "C" courses. This is significantly below the overall mean of 28.1%. Although the NNS students also have a mean significantly below the group mean, when the ESL students are eliminated from that group, the difference is no longer statistically significant.

The logistic regression was then run. The results are presented in Table 14 below. The overall model returned a chi square statistic of 6.13. The probability of a larger chi square statistic by chance alone is about 1.33%. Only the number of "C" courses was significant, but the constant is very close and the number of observations is small.

Table 14: Logit Estimates of Pass 1st by Number of "C" Courses for ESL Students

Pass 1st	Coefficient	Standard Error	Z	P>|Z|	95% Confidence Interval
"C"Courses	.3181964	.1324398	2.403	0.016	.0586192 - .5777737
Constant	-.7069913	.3713679	-1.904	0.057	-1.434859 - .0208765

Once again, the predicted pass rates were obtained, and are reported in Table 15.

Table 15: Predicted Pass Rates for ESL Students by Number of "C" Courses

Courses Completed	0	1	2	3	4	5
Pass Rate	33.0%	40.4%	48.2%	56.2%	63.8%	70.7%

Each course completed raises the probability of passing by about 7%. Using the predicted values of the previous regressions, the ESL students who take no "C" courses can expect their pass rates to trail those of the overall group who took no "C" courses by about 37.8%. However, those ESL students who complete the full sequence would expect their pass rates to lag those of all students who complete the full sequence by only 9.6% and to lag the overall group average of 76.7% by just 6%. While these differences are statistically significant, they suggest that completing the "C" courses was also particularly important for ESL students.

The next difference we wanted to test was for academic preparation. We used the Verbal SAT as a proxy. Many of our students are exempt from submitting SAT scores. Just over 63.5% (N=631) of all freshmen submitted SAT scores including 77.2% of the DSP students, 60% of the ESL students, and 54.2% of the NNS students. 451 of the 631 students who submitted VSAT scores scored below 500. Our first step was to look at the pass rates by VSAT score. To do this we set up 100 point groupings which are reported in Table 16.

Table 16: Observed Pass Rates by VSAT Score

VSAT	200-299	300-399	400-499	500-599	600-699	700 and Above
Pass Rate	51.3%	69.9%	83.4%	86.7%	92.6%	100%
N Size	115	143	193	120	54	6

It seems obvious that whatever the shortcomings of the VSAT, it has fairly strong predictive power for success on the WPR.

I set an arbitrary cutoff of 500 to test for impacts on these students. Among other uses, a score of above 500 on the VSAT exempts incoming freshmen from certain entrance assessment testing mandated by the Massachusetts Board of Higher Education. I then conducted significance tests on the mean number of "C" courses completed by VSAT level. For both of the groups, the

students who failed completed about .4 of a course less than those who passed on the first attempt, which was statistically significant for the below 500 group (|T|=2.3736) but with no statistical significance for the above 500 group. The mean number of completed "C" courses was the same for the two groups.

Table 17: Observed Pass Rates for Students with VSAT Scores by VSAT Level and Number of "C" Course

Courses Completed	0	1	2	3	4	5 or More
>500 Pass Rate	91.7%	83.3%	85.3%	83.3%	90.0%	94.0%
> 500 N Size	24	12	34	30	27	50
<500 Pass Rate	59.1%	72.4%	64.9%	66.7%	75.3%	77.4%
< 500 N Size	44	58	74	63	97	115

Here we see that regardless of the number of "C" courses completed, the mean pass rate for students with above 500 scores is above that of the students with below 500 scores, and above the mean for the overall group.

Logistic regressions were run for the group as a whole and for those above and below 500. The results were significant for the group as a whole, but virtually all of the power is for the students below 500. For the group with 500 or above scores, there is no statistically significant relationship between the number of "C" course and passing the WPR on the first attempt. However, for the group who scored below 500, the relationship was significant. The results are presented in Table 18 below. The overall model returned a chi square statistic of 5.59. The probability of a larger chi square statistic by chance alone is about 1.81%. Both the number of "C" courses and the constant were significant.

Table 18: Logit Estimates of Pass 1st by Number of "C" Courses for Students with Sub-500 VSAT Scores

Pass 1st	Coefficient	Standard Error	Z	P>|Z|	95% Confidence Interval
"C"Courses	.1384777	.0588701	2.352	0.019	.0230945 - .253861
Constant	.4791467	..2001355	2.394	0.017	.0868885 - .871405

Once again, the predicted pass rates were obtained, and are reported in Table 19.

Table 19: Predicted Pass Rates for Students with Sub-500 VSAT Scores by Number Of "C" Courses

Courses Completed	0	1	2	3	4	5
Pass Rate	67.4%	69.6%	71.7%	73.7%	75.6%	77.8%

Each course completed raises the probability of passing by about 2%. Completion of the full sequence of courses raises the probability of passing on the first attempt to above that of the overall group.

Conclusions

The "C" courses were designed to teach a number of skills and to prepare the student for the Writing Proficiency Requirement. Because most students did not fully comply with the "C" course requirement, the variation in the number of courses completed allowed us to assess the effectiveness of the "C" courses in preparing students for the WPR.

It seems clear that there is a positive relationship between the number of "C" courses completed and success on the first WPR attempt. However, it is also clear that this relationship is not equally strong for all students. The effect is strong for ESL students, other non-native English speakers, those who enter through the DSP Program, and those who enter with lower Verbal SAT scores. With the exception of the sub-500 VSAT group, completing the full sequence of "C" courses does not fully eliminate the difference in pass rates between these students and the overall group. However, we would expect that the gap would be closed considerably if the sequence had been completed. When a logistic regression is run for only those students who do not fit into any of these categories (N=466), there is no statistically significant relationship between the number of "C" courses completed and success on the WPR.

Two groups of students completed significantly fewer "C" courses than their comparison groups. The DSP and the ESL students both completed "C" courses at significantly lower rates. It should be noted that these two groups are required to complete a number of courses that do not carry degree credit. It may be that they are less likely that other students to complete other "off track" courses even though the courses may be particularly beneficial for them.

Given the widespread noncompliance with the old system, it is reasonable to believe that there will be similar behavior in the new system. If so, it is important that as resources are dedicated to tracking students and encouraging them to fully utilize the system, these resources and efforts should be focused on these groups of students who most benefited under the old system.

References

Hamp-Lyons, Liz (1996). The Challenges of Second-Language Writing Assessment. In Assessment of Writing: Politics, Policies, Practices. New York: The Modern Language Association of America

Hughes, Gail F. (1996). The Need for Clear Purposes and New Approaches to the Evaluation of Writing-across-the Curriculum Programs. In Assessment of Writing: Politics, Policies, Practices. New York: The Modern Language Association of America

AN ANALYSIS OF MARKET FACTORS, FISCAL STRENGTH, AND INSTITUTIONAL COMPETITIVENESS

Erika M. Newcomer
Research Associate

Jason P. Casey
Director

Higher Education Data Sharing Consortium
Franklin & Marshall College

Abstract

Recent conditions have nourished a competitive environment in which tuition and institutionally-funded discounts spiraled. Using data from 36 private liberal arts colleges, the authors investigated the interrelationship of market demand, institutional wealth, sticker price, and student profile to define an empirical model for estimating stress from increased unfunded tuition discounting.

Introduction

Rising Tuition

The last two decades of the twentieth century were ones of increasingly intensive competition between postsecondary institutions. Economist Caroline Hoxby (in Trusteeship 2000) explained the increased competition as a combination of increased demand for high intensity education, decreased demand for low skilled labor workers, and increased student awareness and mobility. In another work she (Hoxby, 1997) also explained that the market for education became more competitive as it evolved into a nationally and regionally integrated market as students became more mobile. This competition has resulted in a drop in the market share of private institutions as students begin to choose lower-priced public alternatives to private education. At the beginning of the century, 4 out of every 5 students were enrolled in private colleges. The falling private sector market share has resulted in current private institution enrollment of 1 in every 5 students.

Economic analyses have suggested that below-cost tuition at public colleges has drawn students away from the private sector (Wolfram 1997). McPherson (1978) found that the rise in the tuition gap between the public and private sectors accounted for nearly half of the enrollment shift. His regression analysis found the effect of public tuition levels on enrollment to be the strongest at less selective liberal arts colleges. This is extremely important for moderately selective colleges who are adversely affected by institutional density (public institutions per capita) and by low-tuition prices at public counterparts (Thompson and Zumeta 1998).

McPherson and Shapiro (1994) pointed to liberal arts colleges in the Midwest as colleges which specifically faced enrollment declines and severe public sector competition. The same authors (McPherson and Shapiro, 1991) previously suggested that the elite and most-likely well-endowed private colleges provided sufficient financial aid to mute potential enrollment declines resulting from growing discrepancy between their prices and prices at potential public alternatives.

Caroline Hoxby (1997) illustrated how the changing market structure described above explained tuition increases of 50% or more for selective, private colleges. The competitive environment drove institutions to raise their quality (described as vertical differentiation between institutions) in order to improve their market position. Clotfelter (1996) also demonstrated how elite colleges spend money to create "high quality," indicating that an increase in quality cannot occur without subsequent increases in expenses and ultimately, tuition. As Russo and Coomes (2000) pointed out, private colleges are in a precarious position since they are tuition driven and thus must increase tuition to shoulder increased institutional costs.

There also appears to be a relationship between institutional price and reputation of quality. Others (Duffy and Goldberg 1998) point to the "Chivas Regal Effect" which suggests that higher price equals higher superiority. This phenomenon is often cited as the explanation for the actions of many private colleges who significantly raised tuition to emulate the price increases of the Ivy League. Economist Charles Clotfelter (1996) also expressed that a major contributor to rising tuition costs was institutions' "unbounded aspirations" to be "the best" and McPherson and Winston (1993) also acknowledge that high tuition itself may very well be a symbol of quality.

An Environment for Increased Tuition Discounting

Precedents in case law regarding the sharing of enrollment management information also nourished an environment in which sticker prices and institutionally funded discounts spiraled. When the Justice Department ended aid collaboration among the group of 23 highly selective colleges known as the Overlap Group in 1991, it intensified the competition among prestigious institutions. Although Congress did pass legislation which allowed institutions to agree not to engage in non-need merit competition, McPherson and Shapiro (1994) warned that the end of the overlap agreements would eventually cause merit scholarship competition to spread more widely among those elite institutions.

Concurrently, a bull stock market and diversification of services in higher education led to unprecedented financial gains at many, although by no means all institutions. The strong market and the trend toward bolder and more diverse investments resulted in a move of assets from safe bonds into riskier, non-marketable securities pushed endowments up and resulted in several years of double digit returns on institutional investments. Historically, Breneman (2000) explained, annual endowment gains of 9 to 11 percent, coupled with a 4 to 5 percent inflation rate and spending rates of 5 to 6 percent held endowments steady. However, recent year gains

of over 20 percent, lower inflation (2 to 3 percent) and unchanged or even reduced spending have resulted in substantially increased endowment wealth at schools with the largest endowments.

Increased competition and wealth led to radical changes in financial aid policy and practice at many wealthy institutions including shifts in grants from need to merit, replacement of loans with grants, and even tuition freezes. In order to stay marketable for quality students and to bolster their budgets, other institutions have followed suit, mostly by increasing the amount of institutional grant money offered to incoming freshmen. Colleges with a more precarious market position (low selectivity and low yield) began to increase their use of tuition discounting as an enrollment tool to improve their capacity to matriculate students.

Prevalence of Tuition Discounting

How prevalent is the practice of tuition discounting? The 1999 NACUBO Tuition Discounting Survey (as described in Lapovsky and Hubbell 2000) reported an average discount rate (defined by NACUBO as institutional financial aid dollars divided by gross tuition and fee revenue) of 37 percent among independent institutions. Tuition discounting has been on the rise; in the Fall of 1999, 79.4 percent of students received institutional financial aid, compared with 63.7 percent in the Fall of 1990. Additionally, the number of students who do not receive institutional financial aid dropped significantly between that same 10-year period. However, while the number of students receiving aid has increased, the average grant size as a percentage of tuition has remained relatively stable (increasing only 6.5 percent over the same 10-years).

Institutional grants and discount rates vary widely in amounts across colleges. According to researcher Donald Basch (1996), these differences appear related to dissimilarities in selectivity, price, endowment, and location. Compared to the least competitive schools, the most competitive institutions (as characterized by Barron's) were inclined to have lower percentage of needy students, lower average institutional grants, and ultimately, lower discount rates. He ranked colleges by price and noticed that although higher price was associated with a higher discount, the highest discount rates occurred at the 9^{th} and not the uppermost (10^{th}) decile. He also noted that the percentage of needy students declined as price increased (not surprising since evidence, (Zemsky and Oedel, 1983) shows that family's wealth is positively correlated with student's academic qualifications at highly selective colleges); however the needy students at the higher-priced colleges tended to have more need.

Basch (1996) also examined the relationship between endowment and discount, and suggested that a higher endowment allows an institution to provide more generous aid packages. His analyses revealed that higher endowment per student had a statistically significant positive effect on average institutional grant and discount rate.

Problems with Tuition Discounting

This growing trend of tuition discounting may prove stressful for less wealthy institutions, resulting in significant losses of tuition revenue which can be extremely detrimental for schools whose institutional grants are largely unfunded by endowment funds or annual gifts. The National Institute of Independent Colleges and Universities' (NIICU) study revealed that endowment and gift revenue supplied only 30 percent of money necessary to support need-based grants in 1987-88, leaving 70 percent of grants funded by tuition and other revenue sources for most institutions (Evangelauf 1990). As the USA Group Foundation (NACUBO 2000) warns, high discount rates may leave institutions at risk of having fewer dollars accessible for educational expenses.

The USA Group Foundation Study on tuition discounting (NACUBO 2000) found that institutions that offered the largest tuition discounts spent an average of $3,400 per student and lost at least $306 in per-student tuition revenue. A few highly selective institutions lost upwards of $800 per-student tuition dollars and a substantial loss was felt for at least 25 percent of institutions that discounted tuition. Losses in tuition revenue are dangerous as they may potentially lead to losses in academic spending. McPherson and Shapiro (1998) demonstrated that between 1987 and 1994 financial aid increases wiped out a good portion of the private sector's gross tuition increases. Forty-four private liberal arts colleges in their sample increased their sticker prices at an annual real rate of 3.82% while tuition revenues increased at a rate of only 2.7% per year.

Tuition revenue losses may be extremely dangerous to schools without large endowments to help fund discounts, especially if schools with smaller endowments have high discount rates; however schools with smaller endowment funds are just as actively providing institutional discounts. In fact, NACUBO's most recent survey on tuition discounting (Lapovsky and Hubbell 2000) shows that no significant relationship exists between endowment size and level of tuition discount. Actually, there is a slight shift to higher levels of discount as endowment declines. The NACUBO study revealed that the difference in discount rates between the institutions with the largest endowments ($1 billion or more) and those with smaller endowments (less than $50 million) is only 6.2 percent, revealing that endowment size does not determine the size of the institutional discount.

Estimating Stress

This study investigated the interrelationship of market demand (selectivity and yield), institutional wealth (endowment per capita), sticker price, student need and percentage of in-state enrollment. We examined which of these factors predicted higher levels of unfunded institutional grants in an attempt to discern a measure for estimating institutional stress. By looking at predictors of unfunded institutional grants, we may determine potential stress indicators – levels that may alert an institution that it may be nearing a dangerous situation.

Methodology

Data

Data were collected from several sources. General institutional information was extracted from the Integrated Postsecondary Education Data System (IPEDS) Institutional Characteristics Survey of Fall, 2000 (National Center for Education Statistics, 2000). This included institution name, federal id code, state, and locale codes. Year 2000 Carnegie classifications were merged with the remaining general information records from spreadsheets made publicly available by the Carnegie Foundation for the Advancement of Teaching (2000).

Enrollment information was extracted from the IPEDS 2000 Fall Enrollment Survey (National Center for Education Statistics, 2000), including full- and part-time headcounts of degree-seeking students. These counts were used to estimate full-time equivalent (FTE) enrollment for each institution as full-time plus one-third of part-time enrollment. In-state enrollment counts were also derived from the IPEDS Fall Enrollment Survey by dividing the count of first-year students from in-state by the total first-year enrollment.

Demand indices were estimated using data collected on the HEDS Consortium's Freshman Admissions Survey for the fall of 2000. Rejection rates were calculated as the proportion of rejected students to total actionable applicants. Yield rates were estimated as the proportion of matriculating students to the count of admitted students. The demand index, which was used in our statistical models, was estimated as the log of the product of rejection and yield.

Endowment data were obtained from the National Association of College and University Business Officers' (NACUBO) Endowment Study (NACUBO, 2000). The numbers used were end-of-year market value of endowments as of July 1999. These values were divided by FTE for use in our statistical analyses, labeled wealth index, or endowment per capita.

The unfunded institutional grant aid percentage was estimated from IPEDS Financial Survey (F2) data (National Center for Education Statistics, 2000). The proportion of unfunded grant aid was estimated as the proportion of total unfunded student aid dollars to total institutional grant aid dollars.

Aggregate student need, tuition and fees, and total institutional grants were extracted from the Freshman Financial Aid Survey collected by the HEDS Consortium in the fall of 2000. Aggregate need and institutional grants were divided by FTE to estimate a per capita quantity.

Three variables were transformed prior to analysis to make their distributions more normal. Sticker price was squared prior to analysis to correct for negative skewing whereas the square roots of in-state enrollment and endowment per FTE were used to minimize the effects of negative skewing in each of those variables.

Analyses

We analyzed two models using multiple regression analysis: Model I, in which we predicted unfunded institutional grant value by the full predictor set including sticker price, need per FTE, in-state enrollment, endowment per FTE, and demand index. Model II was determined by removing terms from Model I via stepwise regression. For interpretive simplicity, given the disparate metrics of our predictors, we standardized all variables prior to analysis.

Results

Results for multiple regression analyses are summarized in Table 1. Results for Model I (Full Model) were not statistically significant (*Multiple R* = 0.428, $F = 1.301$, $p = 0.291$) although the predictors explained more than 18% of the variance in unfunded institutional grant aid.

Inspection of correlation coefficients (Table 2) revealed (a) a correlation between student need and unfunded institutional aid of 0.34; and (b) large correlations between In-state enrollment and Sticker price ($r = -0.596$), demand index and sticker price ($r = 0.573$), endowment per FTE and in-state enrollment ($r = -0.635$), demand index and in-state enrollment ($r = -0.520$), and endowment per FTE and demand index ($r = 0.300$).

Given the sizable correlations and the likely collinearity among the predictors, we ran a second model (Model II, Table 2) in which we eliminated predictors via backward stepwise progression. This produced a model (Model II in Table 1) which was statistically significant (*Multiple R* = 0.342, $F = 4.513$, $p = 0.041$) in which need per FTE predicted 12% of the variance in unfunded institutional aid [$b = 0.314(0.148)$, $t = 2.124$, $p = .041$].

Figure 1 reveals the relationship of need per FTE and unfunded institutional grant aid. Although the pattern was linear, the regression line did not cleanly fit the observed pattern of the scatterplot. Endowment per FTE, represented by the size of the pips seemed to be unrelated to this relationship.

Because of the expectation of a relationship between sticker price and institutional aid, a follow-up plot of these variables was constructed to determine if the relationship was nonlinear, thus obscuring the relationship (Figure 2). In this plot, a contour was overlaid revealing clustering in the data. Interestingly, there appeared to be two distinct clusters of institutions with parallel linear patterns, suggesting some differentiation of aid strategy based on some other criteria. In this plot, pip sizing based on need per FTE did not suggest that it adequately explained the relationship.

Figure 3 represents the same data as in Figure 2 replotted with pip sizes determined by the demand index. This plot clearly suggests that demand may be a differentiating factor among the two groups of institutions. However, some institutions with similar demand and pricing still appeared to cluster differently based on some unobserved characteristic.

In Figure 4, we replotted the same information with the percentage of students from in-state

as the pip sizing criteria. This sizing parameter suggested another potential predictor of the apparent parallel linear clusters in the data. A comparison of Figures 3 and 4, suggests that the in-state enrollment variable, along with the demand index may adequately explain these apparent clusters.

To test these observations, we ran a final multiple regression analysis in which we used a dummy variable to represent the two clusters of institutions suggested by our model in addition to all previously included predictors (Table 3). This model fit the data much better than the first models, explaining more than 70% of the variance in unfunded institutional grant aid [Multiple $R = 0.840$, $F = 14.841$, $p < .001$). Interestingly, stepwise fitting dropped the only predictor that was included in Model II – need per FTE. It is noteworthy that in-state enrollment and demand Index, while included in the final model specification, were not statistically significant, possibly confirming their importance in creating the two clusters that we observed.

Discussion

Although student need predicts unfunded institutional grant aid as expected, it does not adequately explain the variance in these aid amounts and, by extension, the policies that underlie them. This suggests that as institutions move from largely need-based grant aid to more merit aid as an enrollment tool, factors other than student need will explain the varying amount of grant aid supplied to incoming freshmen.

Viewed simply, the relationship between sticker price and unfunded institutional aid appears to be non-linear. However, when inspected closely, it appears that there are parallel linear patterns for two distinct groups of institutions: one group that has high demand and gives less aid per student and a second group that has less demand and gives larger grants per student. This supports Basch's (1996) finding that although higher price was associated with a higher discount, the highest discount rates occurred at the 9^{th} and not the 10^{th} decile of colleges ranked by price. A more complex situation is setting the amount of unfunded institutional grant if need and price alone do not predict the amount.

Aside from demand, the percentage of students from within state seems to best differentiate the two groups that were observed in the plots of our study. This seems to confirm earlier research that revealed that low tuition at public institutions adversely effected enrollment and tuition prices at private colleges who directly competed with those in-state public schools for enrollment. Private colleges that draw a large base of their students from in-state populations must increase grant amounts to reduce competition from in-state public universities.

Contrary to expectations, there seemed to be little relationship between institutional wealth and grant size. However, given the limited focus of our study on unfunded institutional grant aid as a potential indicator of stress, we necessarily ignored another potential stressor that is likely more closely related to wealth – tuition dependence. That is, an institution with low demand, drawing heavily from within state, giving larger unfunded grants, and heavily dependent on tuition revenues is potentially on shakier fiscal footing than a similar institution that has more

endowment returns with which to pay the bills.

When included as a predictor, the clustering that was observed in the plots improved the overall fit of the predictive model and explained 55% more of the variance than the best model that did not include some indication of the clustering. Although we did not rigorously investigate the factors that determine institutional clustering, our observations of regression results and of the graphics suggests that private colleges that experience moderate to low demand and recruit heavily from within state will likely recycle tuition at greater rates as the sticker price is increased. This suggests that these factors should perhaps be monitored closely, especially for moderately selective institutions.

Future directions

This study was a limited "first look" that was focused on unfunded institutional grant aid as a stress criterion. As was noted above, there are potentially many other factors to investigate as stressors. Most especially, tuition dependence is an important factor that warrants further investigation. Many of the trends in tuition recycling that were examined here become more critical at institutions where tuition makes up a significantly higher proportion of total revenues.

In future studies, we hope to examine the clusters we observed here more closely. Are the clusters real or merely an artifact of the limited variable space defined by the current study? What are the determinants of membership in one cluster or the other? Are there questions of institutional policy – manipulable factors – that can or do determine migration to one pattern of behavior or the other? If all factors are exogenous, what are the key indicators to watch closely in determining one's institution's "footing?"

It is also important to examine the patterns of non-need aid more closely. In this study, we largely ignored this question in an effort to keep our initial model simple – especially in light of our small sample size and limited degrees of freedom. However, the question of a shift from need-based to non-need-based aid becomes more critical to a correct interpretation of the results observed in this study. The elimination of student need from the final regression equation suggests that factors other than federal or institutional need methodology are playing an increased role in disbursement of student aid at less selective institutions, especially as sticker price rises.

Two factors particularly limited this study. First, it was limited to a snapshot of a single year's indicators and second, the dataset was populated by largely moderately to highly selective colleges. A multi-year analysis will permit us to look at direction of change as another dimension of stress. In future studies we intend to expand our dataset to include more moderately-low and lowly selective colleges to provide more variance in the predictors and outcomes.

References

Allan, Ronald G. (1999). *Tuition Discounts, Institutional Student Aid and Scholarship Allowances*. Sallie Mae Education Institute's Monograph Series. (June 1999).

Basch, Donald, L. (1996). Pricing Differences Among Private Colleges: The Impact of Discounting Through College-Funded Grants. *Journal of Student Financial Aid*. 26 (3), 41-58.

Breneman, David W. (2000). Who'll Warm to a Tuition Freeze? *Trusteeship* May/June, 13-17.

Clotfelter, Charles T. (1996). *Buying the Best: Cost Escalation in Elite Higher Education.* Princeton, N.J.: Princeton University Press.

Evangelauf, Jean (1990). Private College Spending on Student Aid Found Up Sharply. *The Chronicle of Higher Education*, May 16, 1990.

Hoxby, Caroline M. (1997). *How the Changing Market Structure of U.S. Higher Education Explains College Tuition.* NBER Working Paper 6323.

Lapovsky, L., and Hubbell, L.L. (2000) Positioning for Competition. *Business Officer*, (March 2000).

Michael McPherson (1978). The Demand for Higher Education, in Breneman, D., and Finn, C. (ed.), *Public Policy and Private Higher Education*,Washington: Brookings Institution.

McPherson, M.S., and Shapiro, M.O. (1998). *The Student Aid Game: Meeting Need and Rewarding Talent in American Higher Education.* Princeton, N.J.: Princeton University Press.

McPherson, M.S. and Shapiro, M.O. (1991). *Keeping College Affordable* Washington: Brookings Institution, 103.

McPherson, M.S. and Winston, G. (1993). The Economics of Cost, Price, and Quality in U.S. Higher Education in McPherson, M.S., Schapiro, M.O., & Winston, G.C, *Paying the Piper: Productivity, Incentives, and Financing in U.S. Higher Education*, Ann Arbor: The University of Michigan Press.

McPherson, M.S. and Shapiro, M.O. (1994). *Merit Aid: Students, Institutions, and Society*. Discussion Paper Number 25. Williams Project on the Economics of Higher Education.

McPherson, M.S., Schapiro, M.O., & Winston, G.C. (1989). Recent Trends in U.S. Higher Education Costs and Prices: The Role of Government Funding. *The American Economic Review* 79 (2), 253-257.

Monks, James, and Ehrenberg, R.G. (1999). *U.S. News and World Report* Rankings: Why do they matter? *Change*, November/December, 43-51.

National Commission on the Cost of Higher Education (HECC). (1998). *Straight Talk about College Costs and Prices*. Washington, D.C.: American Institutes for Research.

Russo, Joseph A., and Coomes, M.D. (2000). Enrollment Management, Institutional Resources, and the Private College. *New Directions for Student Services* 89, 33-47.

Thompson, F., and Zumeta, W.; (1998). *Effects of Key State Policies on Private Colleges and Universities: Sustaining Private-Sector Capacity to Respond to the Higher Education Access Challenge.* Unpublished Manuscript. www.willamette.edu/~fthompso/EER98.html.

Wolfram, Gary. (1997). *The Threat to Independent Education: Public Subsidies and Private Colleges.* CATO Policy Analysis Number 278. August 15, 1997.

Zemsky, R., and Oedel, P. (1983). *The Structure of College Choice*. New York: College Board.

Table 1
Results of Regression Models I and II

Model I: Full Model

R:	0.428
R^2:	0.183

	coeff.	*se*	*t*	*p*
Intercept	0.027	0.156	0.176	0.862
Sticker Price	0.179	0.241	0.746	0.462
Need per FTE	0.252	0.183	1.381	0.178
In-state enrollment	0.336	0.298	1.127	0.269
Endowment per FTE	0.228	0.237	0.960	0.345
Demand Index	0.166	0.224	0.742	0.464

ANOVA

Source	*ss*	*df*	*ms*	*F*	*p*
Regression	5.387	5	1.077	1.301	0.291
Residual	24.007	29	0.828		

Model II: Model Resulting from Stepwise Regression

R:	0.342
R^2:	0.117

	coeff.	*se*	*t*	*p*
Intercept	-0.004	0.146	-0.026	0.979
Need per FTE	0.314	0.148	2.124	0.041

ANOVA

Source	ss	df	ms	F	p
Regression	3.460	1	3.460	4.513	0.041
Residual	26.064	34	0.767		

Table 2
Correlation coefficient matrix and criterion and predictor variables

	(1)	(2)	(3)	(4)	(5)	(6)
Avg. Unfunded Grant (1)	1.000					
Sticker Price (2)	0.132	1.000				
Need per FTE (3)	0.337	0.035	1.000			
In-state enrollment (4)	0.067	-0.596	0.236	1.000		
Endowment per FTE (5)	0.110	0.193	0.044	-0.635	1.000	
Demand Index (6)	0.105	0.573	-0.210	-0.520	0.300	1.000

Table 3
Model III: Adjusted Model Including Clustering

R:	0.840
R^2:	0.705

	coeff.	se	t	p
Intercept	1.315	0.202	6.514	0.000
Cluster Dummy	-2.178	0.294	-7.402	0.000
Sticker Price	0.947	0.145	6.534	0.000
In-state enrollment	0.269	0.176	1.533	0.136
Endowment per FTE	0.297	0.142	2.089	0.045
Demand Index	0.163	0.132	1.243	0.223

ANOVA

Source	SS	df	MS	F	p
Regression	26.004	5	5.201	14.841	0.000
Residual	10.863	31	0.350		

Figure 1.
Relationship of Need to Unfunded Institutional Grant Size.

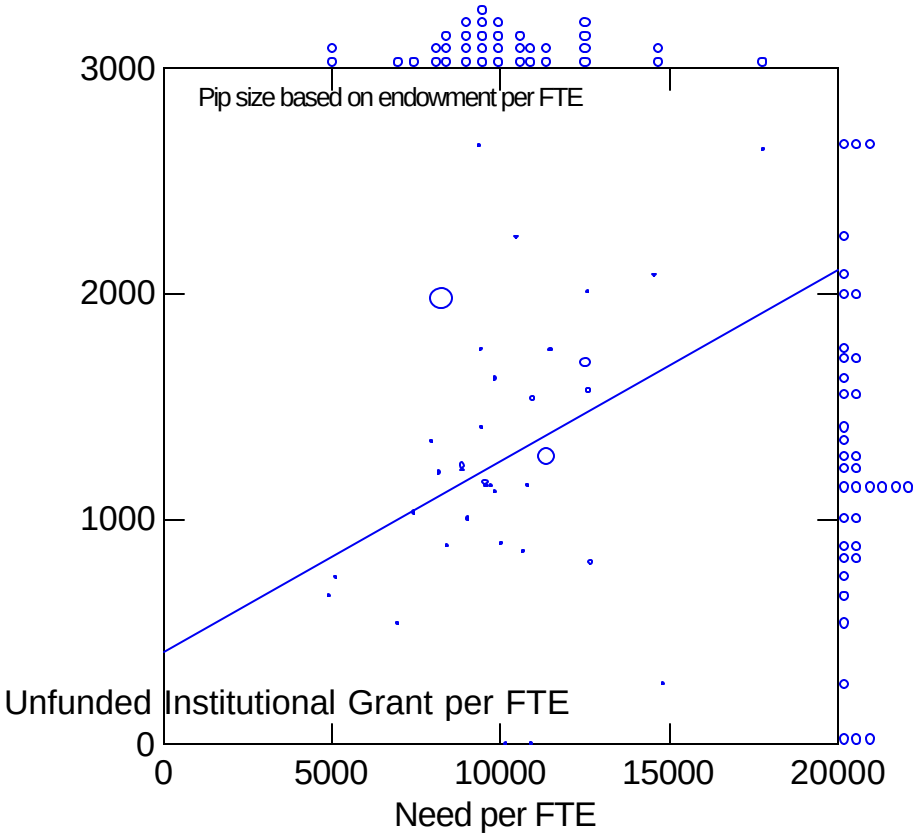


Figure 2.
Relationship of Sticker Price to Unfunded Institutional Grant Size.

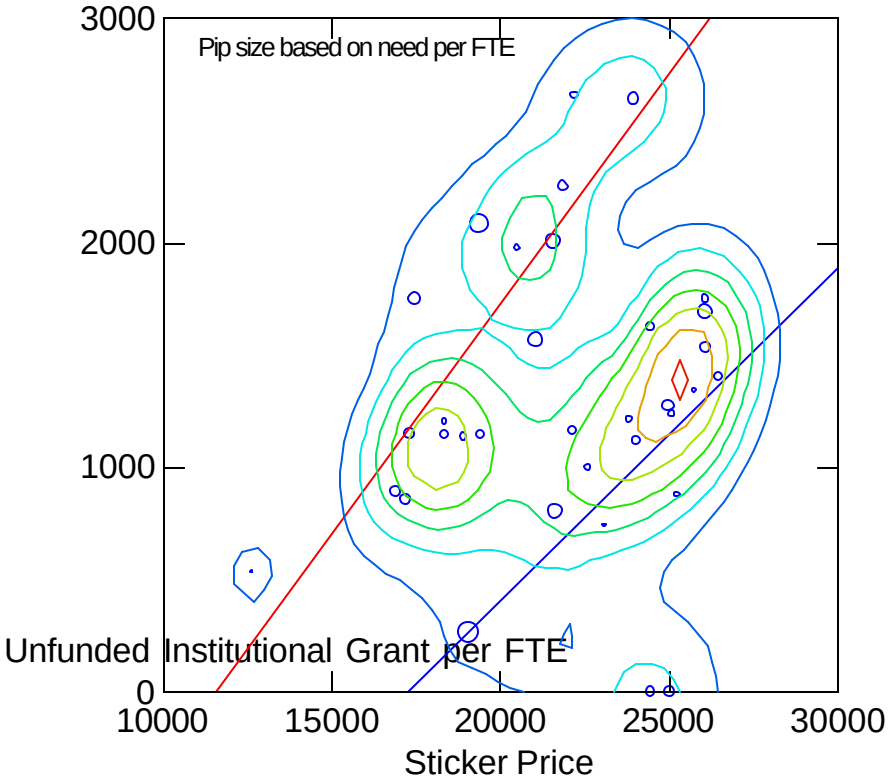


Figure 3.
Relationship of Sticker Price to Unfunded Institutional Grant Size.

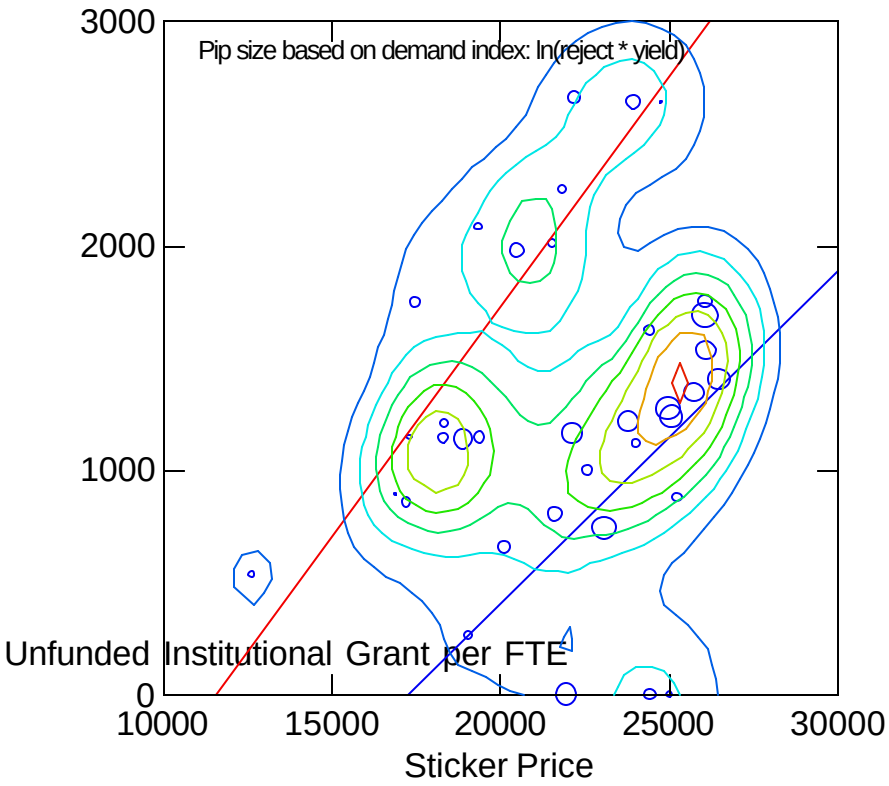
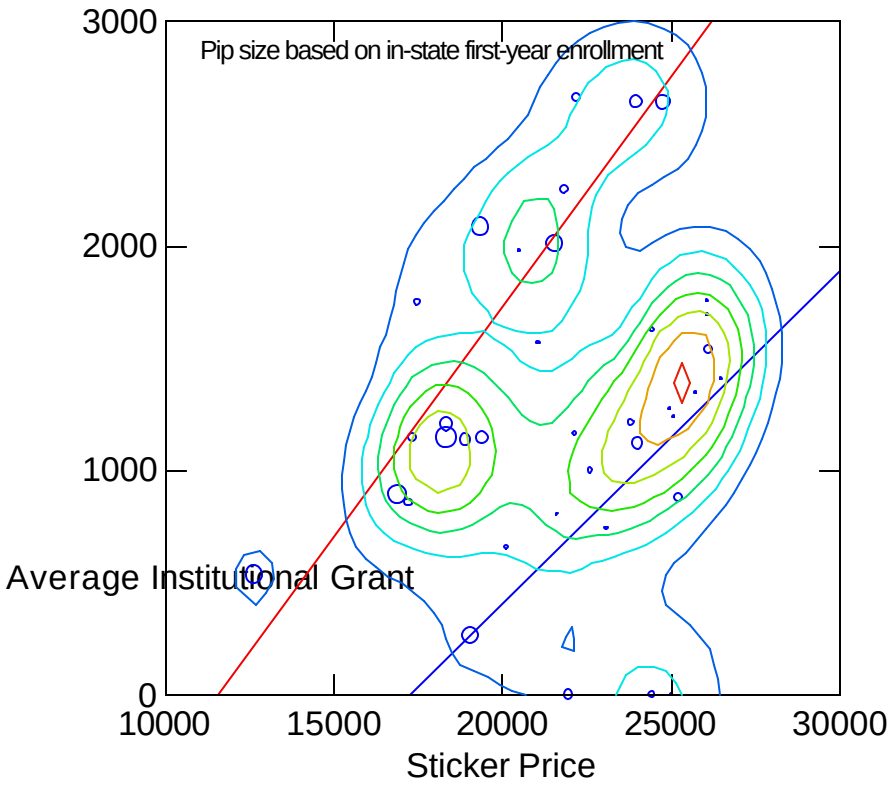


Figure 4.
Relationship of Sticker Price to Unfunded Institutional Grant Size.



THE IMPACT OF INSTRUCTIONAL DELIVERY ON LEARNING OUTCOMES AND INTENT TO PERSIST

Kristin A. Owens
University of Maryland

J. Fredericks Volkwein
Pennsylvania State University

Introduction and Problem Statement

The academic disciplines of criminal justice and higher education have provided little evidence to document the effects of educational participation during inmate incarceration (Almond, 1989; Boaz, 1976; Cheatwood, 1988; Wolf & Sylves, 1981). Research has been hindered by changing societal beliefs about the rehabilitative potential of incarceration, as well as by political and program funding decisions. Pell Grant eligibility and funding for inmates was eliminated in 1994 with the passage of the Violent Crime Control Act. The lack of valid, rigorous research contributed to the downfall of programs since there was little hard evidence in favor of maintaining their existence.

Nevertheless, the enormous growth in the American prison population and the growing importance of educational opportunity and attainment for both individuals and society gives some weight to the need for such research. The theoretical purposes of both traditional higher education and corrections are complimentary; higher education attempts to provide an environment for gaining new knowledge, skills and educational advancement, while corrections applies the principal that attitudes, ideas and behavior can be changed through rehabilitation (Gehring, et al, 1998). However, countless research studies of effectiveness in traditional higher education continue to be produced while college level programs in correctional facilities lack proper evaluation. Few have attempted to apply the current student outcomes models to inmate learners, and decisions about funding for post secondary programs in prisons are often made in the absence of alternatives supported by evidence.

Conceptual Frameworks

In designing this study, we drew upon appropriate frameworks from the scholarly literature. The most traditional view is that pre-college characteristics like student family backgrounds, academic preparedness for college, and clear goals are the main factors accounting for differences in academic performance, persistence behavior, and other

educational outcomes (Feldman & Newcomb, 1969; Astin, 1991; Stark et al. 1989; Willingham et al. 1985).

A second group of alternative yet complementary perspectives fall under the general description of **student-institution fit models** (Pascarella & Terenzini, 1991). These models generally suggest that student persistence and growth depends on the degree of successful integration into the academic and social structures of the institution, and on the amount of student involvement and effort.

A third relevant framework for this study derives from the literature on **self-efficacy** -- a person's judgment of their capabilities to act in order to attain their desired goal or performance (Bandura 1986). Self-efficacy affects choice of activities, goal formulation, effort and persistence to degree attainment (Bandura, 1977; Lent, Brown & Larkin, 1984; Schunk, 1991). Research on inmates has found that the longer a student is in prison, the lower the self-efficacy scores (Parker, 1990).

Based upon these three branches of the scholarly literature, we assume that inmate educational outcomes and reported gains and intent to persist are BI-products of the following factors: demographic backgrounds (including age, sex, and ethnicity), length of prison sentence, instructional method (traditional classroom versus distance), and learning context (faculty effectiveness in the classroom, peer interaction, and learning environment). This research focuses on the role of instructional method, controlling for these other potential influences.

Research Design and Methodology

This study utilizes a cross-sectional research design, collecting information at a single point in time. The subjects of this study are 274 inmates (out of 279 enrolled) from nine correctional institutions in a single state. The survey was administered to 111 students in traditional courses and 163 students in distance education courses. The survey was kept confidential and completely voluntary. The inmates completed the questionnaire within twenty minutes. Several factors prevented the study from following a longitudinal design, including restrictions on information access, confidentiality, and the extreme transience of the prison population due to probation and parole, prison transfers, and programming re-assignments

Based on the outcomes and self-efficacy literature, 44 survey items were designed in the areas listed in Figure 1. In each category, item construction was grounded in theory and research to ensure construct validity. To ensure face or content validity, we consulted experts in higher education and criminal justice, and used a practitioner focus group to assist with item formulation. Survey items were developed, pilot tested, and subsequently revised.

Demographic Information includes age, race, gender, length of sentence, and work hours.

Instructional Method- Students enrolled in distance education courses were coded '1', traditional classroom courses coded '0'.

Faculty Effectiveness- This six-item scale measures the nature of classroom interaction between inmates and their instructors. These six items are drawn from the classroom involvement scale developed by Terenzini, et al (1982, 1984) and adapted for this study (alpha= .89).

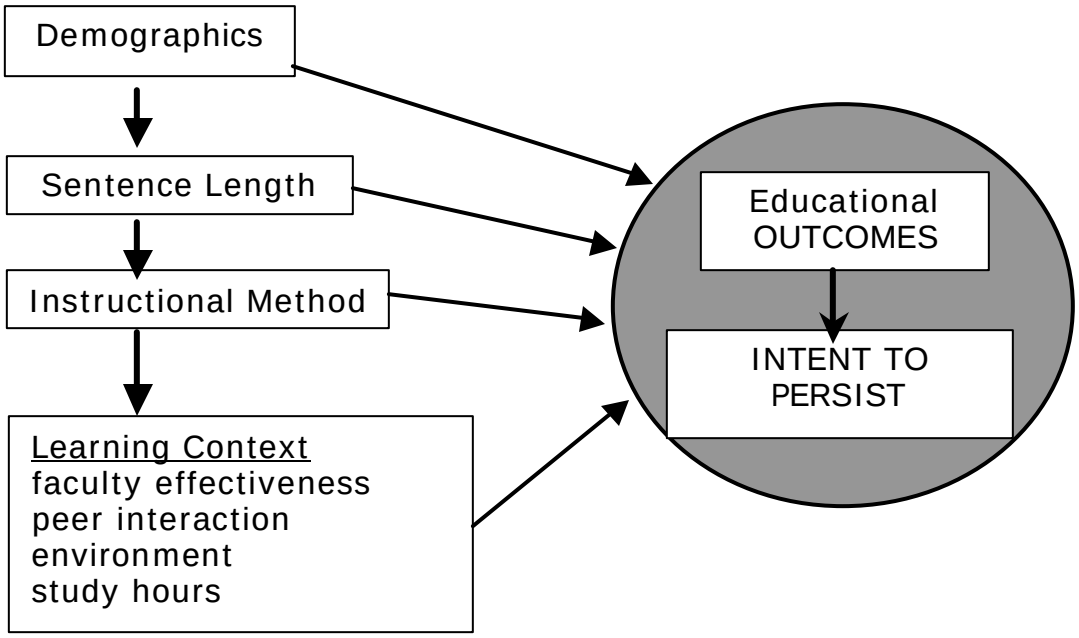
Peer Interaction- This five-item scale measures the nature of the interactions between inmates and their student peers. These five items are drawn from the peers sub-scale of the Mattering Scales For Adult Students in Post-secondary Education (Schlossberg et al., 1990) and modified to meet the needs of this study (alpha=.83).

Learning Environment- This scale is measured by five items describing the learning environment, drawn from the Learning Environment Inventory (Fraser, Anderson, Walberg, 1982), originally designed to assess the classroom learning climate of secondary students, and later adapted to successfully assess the learning environment of college classrooms (Ellet, 1976; Kent & Fisher 1997). The five items were adapted to meet the needs of this study (alpha=.69).

Educational Outcomes- As suggested by Figure 1, a scale of educational outcomes is treated first as a dependent variable, then as an independent predictor of persistence. The outcomes scale is measured by six domains describing career preparation, gains in job skills, problem solving, openness to new ideas, control, and civil responsibility. These six items are drawn from the Noel-Levitz College Student Inventory (Stratil & Schreiner, 1993) and modified to meet the needs of this study (alpha=.86).

Intent to Persist – Two items based on Bandura's concept of self-efficacy, were used to measure intent to persist in the course. Students rated their confidence in completing the course and semester on a 1 to 5 scale (alpha=.76). Three items also based on Bandura's concept of self-efficacy, were used to measure intent to persist towards a degree. Students rated their confidence in persisting to degree attainment on a 1 to 5 scale (alpha=.92).

Figure 1: Research Design Model



Results for Educational Outcomes

We analyzed the relationships among the variables using OLS regression and entered the variables in blocks with listwise deletion of cases. The results are shown in Table 1 and summarized below. The final equation (Model 4) is robust and accounts for almost 36% of the variance in perceived outcomes. The findings from this highly controlled population of prison inmates look very similar to the results we typically have found in outcomes studies of traditional college students.

- *Being female is positively associated with Educational Outcomes.*
- *Faculty classroom effectiveness is highly influential and positive.*
- *Peer Interaction is highly influential and positive.*
- *Inmates receiving distance instruction report significantly lower outcomes, but the influence is indirect.* This particular type of distance instruction exerts a direct negative influence on the measures of the Learning Context, especially Faculty Effectiveness and Peer Interaction, and they in turn influence Outcomes.
- Ethnicity, age, length of prison sentence, and hours of work and study are not significant.

Table 1: OLS Regression Results for Educational Outcomes

Variable Block	Model 1 Std. Beta	Model 2 Std. Beta	Model 3 Std. Beta	Model 4 Std. Beta
1. Demographics:				
Female	**.229****	**.228****	**.136***	**.113***
Minority	-.022	-.022	.003	-.041
Age	-.026	-.025	.050	.007
Work Hours	-.077	-.076	-.036	-.072
2. Length of Sentence		-.006	-.006	-.067
3. Distance Education			**-.252****	-.066
4. Learning Context:				
Faculty Effectiveness				**.262****
Peer Interaction				**.379****
Learning Environment				.066
Study Hours				.002
Total R2	**.052****	**.052****	**.101****	**.357****

* = p<.05 ** = p<.01

Results for Intent to Persist in Course

We analyzed the relationships among the variables using OLS regression and entered the variables in blocks with listwise deletion of cases. The results are shown in Table 2 and summarized below. The final equation (Model 5) is robust and accounts for 37% of the variance in intent to persist in the course/semester.

- Gender, Ethnicity, Age, Sentence Length, and hours of work and study are not significant
- *Distance education students not only report lower Educational Outcomes also report significantly lower Intent to Persist in Courses.* This particular type of distance instruction interacts negatively with the measures of Learning Context, especially Faculty Effectiveness, but also Peer Interaction.
- *Faculty Effectiveness is highly influential and positive –twice as influential as learning outcomes and three times as influential as mode of instruction.*
- As we saw in Table 1, *Peer Interaction* exerts a direct positive influence on the outcomes measure, but in Table 2 peer influence on Intent to Persist in Course is indirect, as its influence disappears once the Outcomes measure is added in Model 5.
- *There is a strong positive statistical connection between perceived*

Outcomes and Intent to Persist in the Course/semester.

Table 2: OLS Regression Results for Intent to Persist in Course/Semester

Variable Block	Model 1 Std. Beta	Model 2 Std. Beta	Model 3 Std. Beta	Model 4 Std. Beta	Model 5 Std. Beta
1. Demographics:					
Female	.103	.103	.080	.080	.055
Minority	-.042	-.042	-.035	-.031	-.022
Age	-.014	-.014	.005	-.084	-.086
Work Hours	.058	.058	.067	.012	.028
2. Length of Sentence		.000	.000	-.044	-.029
3. Distance Education			-.064	**.154****	**.168****
4. Learning Context:					
Faculty Effectiveness				**.528****	**.470****
Peer Interaction				**.183****	.100
Learning Environment				-.009	-.023
Study Hours				.082	.081
5. Outcomes					**.219****
Total R2	**.017**	**.017**	**.021**	**.339****	**.370****

* = p<.05 ** = p<.01

Results for Intent to Persist in Degree

We analyzed the relationships among the variables using OLS regression and entered the variables in blocks with listwise deletion of cases. The results are shown in Table 3 and summarized below. The final equation for degree persistence (Model 5) is not as robust as for outcomes and course persistence, and accounts for 15% of the variance.

- *Age is significant.* Results show that younger students (30 years and younger) are more likely to indicate positive Intent to Persist in Degree than older adults (older than 30 years)
- Being *female* directly influences Intent to Persist in Degree through Model 3. Once Learning Context is added to the model, Gender is no longer significant.
- Ethnicity, Work Hours and Sentence Length are not significant.
- *Distance education has no effect* on Intent to Persist in Degree.
- Results indicate inmates who experience *gains in Educational Outcomes* are also likely to report higher Intent to Persist towards a Degree, controlling for other

factors.

Table 3: OLS Regression Results for Intent to Persist in Degree

Variable Block	Model 1 Std. Beta	Model 2 Std. Beta	Model 3 Std. Beta	Model 4 Std. Beta	Model 5 Std. Beta
1. Demographics:					
Female	**.156***	**.154***	**.144***	.130	.098
Minority	.104	.104	.107	.077	.089
Age	**-.146***	**-.144***	**-.136***	**-.155***	**-.157***
Work Hours	-.012	-.010	-.006	-.024	-.004
2. Length of Sentence		-.009	-.009	-.012	-.007
3. Distance Education			-.026	.026	.045
4. Learning Context:					
Faculty Effectiveness				.013	-.059
Peer Interaction				**.202****	.097
Learning Environment				.021	.003
Study Hours				.104	.104
5. Outcomes					**.277****
Total R2	**.047***	**.047***	**.048***	**.104****	**.154****

* = p<.05 ** = p<.01

Conclusions

The phrase "culture of failures" (Roundtree et al, 1982: 17) is often used when describing the inmate population; however, "if an educational approach can help modify this sense of failure, such an approach deserves attention." Since self is an important determinant of personal adjustment among minority college students, especially inmates, sources of positive self-expectations can be modeled in higher education programming within correctional institutions (Solberg & Villarreal, 1997). The findings of this study should encourage change in existing correctional education program structures, by providing more opportunities for faculty and peer interactions.

There is a belief in the correctional community that educational programs are wasted on long-term prisoners and that they are not motivated to learn. This study finds that sentence length is not significantly associated with student outcomes or intent to persist. As shown in the statistical analysis, inmates with varying sentence lengths reported similar gains in outcomes and intent to persist in both course and degree. Admission criteria for inmate access to college programs should reflect these findings. Program participation should not be limited by sentence length, as it is now in most states. Inmates incarcerated for longer periods of time should not be denied post-secondary educational opportunities just because they will not be released anytime

soon.

Another important finding in the study relates to gender. As a whole, there are far fewer female correctional facilities across the nation. Appropriate staff, learning environments, and materials may be disproportionate when compared to male facilities. Since this study indicates that females are associated with positive outcomes and intent to persist, college programs should be encouraged to expand for women audiences within the correctional system. All incarcerated individuals should be afforded the same level of education experiences, regardless of gender.

Our research suggests that traditional classroom instruction is superior to video instruction for student inmates. For reasons of both budget constraints and security, correctional distance education contributes to rely substantially on correspondence courses and videotapes. Our findings indicate that videotape delivery, as distinct from face-to-face instruction, results in lower educational outcomes and lower intent to persist. This finding is consistent with what we know about the value of active versus passive learning. While we are skeptical about the effectiveness of distance education for this population, we are not necessarily prepared from this study to draw conclusions about distance education using more advanced technologies and synchronous learning. Method of instruction has never before been examined as an indicator of inmate educational outcomes, nor as a contributor to low recidivism. Thus, more research is needed on prison culture, instructional method, and inmate learning.

An important conclusion determined by the literature review and this study is that post-secondary educational outcomes, result from multiple influences (Wolf & Sylves, 1981). Studies on traditional college campuses and prisons alike have shown that multiple variables contribute to student success, learning, and persistence. Further research is needed with inmate students to draw the connection between education and reduced recidivism.

In summary, this study examines the relationships among instructional delivery method, student outcomes, and intent to persist in both course and degree for a population of 274 inmates participating in college programs at nine Maryland State prisons. Consistent with the existing higher education research and literature, faculty classroom effectiveness and inmate peer interactions exert the strongest positive and most direct influences on educational outcomes; and these outcomes in turn have the strongest impact on persistence. Video delivery, as distinct from traditional face-to-face instruction, has a direct negative effect on faculty and peer interactions, and an indirect negative effect on educational outcomes, and on intent to persist. Being female is positively associated with higher outcomes and intent to persist, encouraging the continued and expanded college programming opportunities for women inmates. An unexpected finding is the insignificance of sentence length on both student outcomes and intent to persist, thus supporting more flexible admission criteria for post-secondary correctional programs. Finally, the measures of ethnicity, work hours, and study hours are not influential predictors.

With the continued growth of technology and new distance education avenues, the opportunities for new and enhanced programs have increased. Educators and program

administrators need not answer how can we *teach* utilizing distance instruction technology, but instead, how can we *maximize* student learning (Champagne, 1998: 90). Hopefully this study will enable more positive changes and informed decision-making opportunities for practitioners in the field of correctional education and encourage additional areas of research for academics engaged in student development theory.

References

Almond, K. L. (1989). Adult post-secondary education in secure federal correctional institutions in British Columbia: a comparative survey of prison-students' perceptions (Doctoral dissertation, Simon Fraser University, 1989). Dissertation Abstracts International No. AAT MM59271.

Astin, A.W. (1991). Assessment for excellence. New York, NY: MacMillan.

Bandura, A. (1977). Self-efficacy: the exercise of control. New York, NY: W. H. Freeman and Company.

Boaz, M. E. (1976). An evaluative study of project outreach to inmates: a higher education program offered by the University of Virginia at three of Virginia's correctional institutions (Doctoral dissertation, Virginia Polytechnic Institute and State University, 1976). Dissertation Abstracts International No. AAT 9769577.

Champagne, M. V. (1998). Dynamic evaluation of distance education courses. Proceedings of the Annual Conference on Distance Teaching & Learning. Madison, WI: 89-96.

Cheatwood, D. (1988, December). The impact of the prison environment on the incarcerated learner. Journal of Correctional Education, 39, 184-186.

Ellet, C. D., et al. (1976). The Learning Environment Inventory: a reexamination of its structure and use. Paper presented at the Annual Meeting of the American Educational Research Association. San Francisco, CA. (ERIC Document Reproduction Service No. ED 129 882.

Feldman, K. and Newcomb, T. (1969). The Impact of College on Students. San Francisco, CA: Jossey-Bass.

Fraser, B. J., Anderson, G. J. & Walberg, H. J. (1982). Assessments of learning environments: manual for Learning Environment Inventory (LIE) and My Class Inventory (MCI). Perth, Western Australia: Western Australian Institute of Technology.

Gehring, T., McShane, M. & Eggleston, C. (1998). Then and now: approaches to correctional education in the United States. In W. Forster (Ed.) Education behind bars - international comparisons (pp. 147-166). Leicester, Great Britain: National Institute of Adult Correctional Education.

Kent, H. & Fisher, D. (1997). Associations between teacher personality and classroom environment. Paper presented at the Annual Meeting of the American Educational Research Association. Chicago, IL. (ERIC Document Reproduction Service No. ED 407 395).

Lent, R. W., Brown, S. D. & Larkin, K. C. (1984, July). Relation of self-efficacy expectations to academic achievement and persistence. Journal of Counseling Psychology, 31 (3), 356-362.

Parker, E. A. (1990, September). The social-psychological impact of a college education on the prison inmate. Journal of Correctional Education, 41, (3) 140-146.

Pascarella, E. T. & Terenzini, P. T. (1991). How college affects students. San Francisco, CA: Jossey-Bass.

Roundtree, G. A., et al. (1982, January). The Effects of education on self-esteem of male prison inmates. Journal of Correctional Education, 32 (4), 12-18.

Schlossberg, N. K., Lassalle, A. D. & Golec, R. R. (1990). The Mattering Scales for Adult Students in Postsecondary Education. American Council on Education: Washington, D.C.

Schunk, D. H. (1991). Self-efficacy and academic motivation. Educational Psychologist, 26 (3/4), 207-231.

Solberg, V. S. & Villarreal, P. (1997). Examination of self-efficacy, social support, and stress as predictors of psychological and physical distress among Hispanic college students. Hispanic Journal of Behavioral Sciences, 19 (2), 182-201.

Stark, J. et al. (1989). Student goals for college and courses: a missing link in assessing and improving academic achievement. ASHE-ERIC Higher Education Report 6. (ERIC Documentation Reproduction Service No. ED 317 121)

Stratil, M. & Schreiner, L. (1993). Technical Guide in Retention Management Systems Coordinators Manual. Noel-Levitz Center Inc.: Iowa City, IO.

Terenzini, P. T., Pascarella, E. T., and Lorang, W. (1982). An assessment of the academic and social influences on freshmen year educational outcomes. Review of Higher Education, 5: 86-110.

Terenzini, P. T., Theophilides, C., and Lorang, W. (1984). Influences on students' perception of their personal development during the first three years of college. Research in Higher Education, 21, 178-194.

Willingham, W., Young, J., and Morris, M (1985). Success in college: The role of personal qualities and academic ability. New York: College Entrance Examination Board.

Wolf, J. G. & Sylves, D. (1981). The impact of higher education opportunity programs: post prison experience of disadvantaged students -- a preliminary follow-up of HEOP ex-offenders. Albany, NY: New York State Education Department. (ERIC Documentation Reproduction Service No. ED 226 073)

THE IMPACT OF LOTTERY INCENTIVES ON STUDENT SURVEY RESPONSE RATES

Stephen R. Porter, Director of Institutional Research

Michael E. Whitcomb, Assistant Director of Institutional Research

Wesleyan University

Abstract

Lottery incentives are widely used by institutional researchers despite a lack of research documenting their effectiveness. A controlled experiment tested the effects of lottery incentives using a prospective college applicant web survey, with emails sent to over 9,000 high school students. The impact of the level of lottery incentive on response rates, time to response, and response bias is discussed.

Introduction

Student survey data have grown increasingly important for institutions of higher education. But as the use of student and alumni data have increased, response rates to surveys have been falling nationally (Dey, 1997; Smith, 1995; Steeh, 1981). Survey fatigue is commonly cited, as public opinion polls have become more popular with the media and telemarketers use surveys for data mining research. Increasingly educational researchers are faced with the prospect of simply maintaining, rather than increasing, survey response rates.

As response rates continue to shrink, researchers face increasing costs to counter survey non-response. Second and third mailings, for example, must be larger if the initial mailing elicits a weak response pool. Given that survey research is one of the most common activities in institutional research (Schlitz, 1988), researchers must refine their data collection tools to counter this growing trend.

Based on a survey of colleagues, lottery incentives appear to be a popular method for increasing response rates in institutional research surveys. A lottery incentive is a reward offered to survey recipients for responding to a survey, in which every recipient who responds is entered into a drawing (similar to a lottery) for one or more prizes.

In addition, the growing use of electronic surveys may be leading to an increased use of lottery incentives in survey research. Unlike mailed surveys, it is impossible to include incentives such as a dollar bill with an email survey or email notice about a web survey (Couper, 2000). Incentives paid upon completion are possible with electronic surveys, however, as these can always be mailed to respondents. With the growing use of web surveys the need to understand the efficacy of lottery incentives has increased.

While lottery incentives appear to be a popular and perhaps growing method for increasing response rates, the research literature on lottery incentives indicates they have little or

no impact on survey response. When applied to higher education, however, this research may not be relevant. Previous studies have been conducted on members of the general population, and it may be possible that college students are more price-sensitive than the average person. If so, lottery incentives may have an impact on response rates in student surveys while not having any impact on surveys of the general population.

Educational researchers have investigated the impact of questionnaire format (Boser, 1990), survey length (Adams & Gale, 1982), bad addresses (Grosset, 1995) and multiple follow-ups (Cote et al., 1986; Smith & Bers, 1987) on student and former student survey response, but no research has been conducted studying the impact of lottery incentives on response rates. The lack of research in this area is not surprising. To truly test the effectiveness of incentives, at least two randomly selected groups of students must be used: a control group receiving just a survey, and one or more experimental groups receiving both a survey and an incentive. Comparing response rates across different administrations of the survey (e.g., across years) does not work, as other factors may affect response rates besides a change in survey administration. A similar rationale holds for surveys across institutions within a university system.

Yet such an experimental situation poses a problem for any college or university, as students in the control group will discover that other students have the possibility of winning a prize for filling out the same survey. Given communication among students and the likely negative reaction of the control group, successful implementation of a controlled experiment of incentives is simply not possible on our campuses.

We circumvent this problem by conducting a controlled experiment on high school students who have contacted an admissions office at a small, liberal arts college for information about applying. Since this college draws students from across the nation, we can split the survey into groups without worrying about communication between groups. And because the high school students are seniors about to attend college, their sensitivity to a lottery incentive should be much more similar to the typical college student's sensitivity compared with a member of the general population.

Our research is useful given the prevalent use of lottery incentives in institutional research, the lack of evidence of their efficacy, and the probable increase in their use as researchers begin moving from paper to electronic surveys. The paper consists of four parts. After first reviewing the literature on lottery incentives and response rates, we determine the prevalence of such incentives in institutional research. We next describe the experiment and how it was conducted. We then analyze the data for differences in response rates and time to response. We also examine item non-response differences between the experimental groups.

Literature review

Incentives are theorized to affect response rates by affecting how the respondent views the costs and benefits of the survey process. Social exchange theory proponents argue that by

providing general, diffuse, or "token" benefits to potential respondents, researchers can convince potential respondents "that in the long run the anticipated benefits of responding outweigh the anticipated costs," (Dillman et al, 1996, p. 378). Such benefits can include small amounts of money (usually in the $1-$5 range), mention of benefits that will accrue to groups to whom the respondent belongs, assistance to the research sponsor, or benefits to the population as a whole (Dillman et al., 1996).

Economic exchange theory, on the other hand, posits that respondents will fill out and return surveys in exchange for specific monetary compensation, rather than a broader, more diffuse benefit. Dillman (2000, pp.14-15) argues that economic exchange simply does not work with surveys. As evidence he cites survey research that promised payment to respondents upon completion, and which found no increase in response rates using this method.

If Dillman were correct, then we would expect lottery incentives to have no effect on response rates. The issue here is that the payment of the incentive after the survey has been returned turns the survey process into an economic transaction rather than a social one. Prepaid incentives work precisely because they are obtained without any effort on the part of the respondent, and because they may create a sense of duty on the part of the respondent. This norm of reciprocity (Groves et al., 1996) arises because the token incentives are viewed as a gift rather than compensation for effort.

Alternatively, an economist might argue that lottery incentives do not work simply because the benefits appear too diffuse. With a lottery the expected benefit is not the monetary amount of the incentive, but the amount multiplied by the probability that the respondent will be selected a winner in the lottery. The implication is that larger lottery incentives might have an impact, as they will have a larger expected value for the respondent. (Alternatively, increasing the probability of winning should also have the same effect.) An additional complication here is to what extent the respondent actually believes that a lottery exists and will be run fairly.

The empirical research on incentives indicates a conclusive positive impact on response rates; however, this impact very much depends on the type of incentive. Incentives can be divided into two groups based on when the survey recipient receives the incentive: either with the survey (known as pre-payment) or after the survey has been completed and returned (post-payment).

Numerous studies have been conducted studying the impact of pre-paid incentives on survey response, and the results indicate that their use almost invariably increases response rates (e.g., Church, 1993; Singer et al., 1999; Willimack et al., 1995; Zusman and Duby, 1987). Less certain is the impact of post-paid incentives. Several experimental studies have been conducted that compare the impact of pre- and post-payment of incentives, with the general finding that post-payments have no statistically significant impact on response rates (Church, 1993; James & Bolstein, 1992; Singer at al., 2000). Other researchers have tested the effect of lottery post-payments, in which the incentive is not guaranteed but is instead dependent upon the outcome of a drawing. These researchers have found no effect for lottery incentives (Warriner et al., 1996).

The use of lottery incentives in institutional research

Given a lack of theoretical and empirical support for the use of lottery incentives in surveys, the extent of their use in institutional research is surprising. In Spring 2000 we conducted a short web survey asking institutional researchers about their use of lottery incentives. Members of seven regional institutional research listservs (California, Mid-America, Northeast, Pacific Northwest, Rocky Mountain, Southern, Upper Midwest) were notified of the survey, with 374 people responding. Respondents were asked several questions about the surveys they conduct, and whether their institution was public or private.

Table 1 shows the number of surveys in which researchers used a lottery incentive during a typical academic year. Overall about a third of the respondents administer at least one survey a year that uses a lottery incentive, with about half of the respondents at private institutions using lottery incentives at least once a year.

Researchers were also asked what types of incentives they use. Table 2 lists the type of prize. Monetary prizes (either cash or gift certificates) tend to predominate, with respondents at private institutions more likely to list these as the incentives they use. The differences between private and public institutions seen in Tables 1 and 2 most likely stem from differential resources, with private institutions using monetary incentives and using them more often than their public counterparts.

Researchers were also asked their opinion of the effect of lottery incentives on response rates. Table lists the responses for all respondents, and only those reporting that they used a lottery incentive in at least one survey per year. Given the prevalent usage seen in Table 1, not surprisingly 75% of respondents reported that they believed lottery incentives increase response rates, with the vast majority indicating that they 'somewhat' rather than 'greatly' increase response rates. About 90% of respondents who indicated they use lottery incentives in their surveys believe incentives positively increase response rates, while over half of those who do not use incentives still believe in their efficacy.

In sum, although the literature on incentives and response rates shows that post-payment of incentives in general and lotteries in particular have little or no impact on survey response, use of such lottery incentives appears common in institutional research. The remainder of the paper investigates whether lottery incentives are indeed effective when used with student surveys.

Research design

The experiment was conducted in Spring 2001 during a survey of non-applicant high school students. These prospective students had contacted the institution for information about the institution during the previous year, but did not apply for admission. Of about 13,000 prospects, 9,305 had provided enough information about their high school during the contact to allow the assignation of the appropriate CEEB code for their high school. Because it was essential that members of the control group did not discover that other students had been offered an incentive for

response, students were grouped by high school for the experiment. The average number of students per high school was 2.64, with the number of students ranging from 1 to 93.

The high school codes were randomly divided into five groups: a control group and four incentive groups. The high schools codes were then used to assign students to an experimental group. This ensured that students in the same high school were placed into the same experimental group, and therefore would not discover via communication with friends that others in their high school had received a different incentive offer. Table 4 shows the number of high schools and number of students in each experimental group.

The survey administration consisted of an initial email notification with an embedded survey link, and each group was administered the same survey. Although students were asked to enter their email address during the survey, the sample groups were given links to five separate websites to ensure we could track differences between groups. The four randomly selected incentive groups were informed that if they responded to the survey they would be entered into a drawing for a $50, $100, $150 or $200 gift certificate to Amazon.com, depending on the group. The emails were identical except for the incentive group emails, which included this passage about the lottery incentive:

Because we realize your time is valuable, when you complete the survey you will be entered into a drawing for a $__ gift certificate from Amazon.com. The drawing will be held within six weeks and you will be notified of the outcome via email.

The initial email was followed three days later with a reminder email to non-respondents, with a final reminder to non-respondents five days after the first reminder. Each reminder included details about the incentive for each group.

Results

In all analyses we examined the control and experimental groups to test three main questions:

- Do all five groups (the control and four levels of incentive) differ from one another?
- Does each incentive level individually differ from the control group?
- Do respondents offered an incentive differ from those not offered and incentive?

The first question tests if increasing levels of incentives have a differential impact, in other words, do response rates increase as the amount of the incentive increases? This is the most common view of incentives and their impact on response rates: more is better.

The second question tests if only some of the incentives have an impact. For example, there may be a nonlinear relationship between response rates and incentives amounts. Small amounts may have little impact because the respondent does not feel they are adequate to justify his or her expenditure of time. Large amounts, on the other hand, may have little impact because respondents are skeptical they will receive the prize given the large value.

Alternatively, large amounts may be viewed as compensation rather than a token benefit, thus transforming the relationship between one of reciprocity to an economic one.

The third question simply tests the overall impact of offering an incentive. There may not be much of a difference in response rates between the $50 group and the $200 group, and depending on the data, ANOVA testing for differences between all five groups could result in a null finding. Yet an ANOVA testing the control versus all the incentive groups might detect a positive impact, so this third hypothesis is simply another way to check the data.

Table 5 shows the response rates for the initial email and the response rates at the close of the experiment. Rates are shown for all five groups, for the four incentive groups combined, and for the entire sample. Overall 15.2% of the sample responded to the survey. Differences between the control group and incentive groups were quite small. Almost 14 percent of the control group responded, while overall 15.6% of respondents in the incentive groups participated in the survey.

To test for differences in response rates both following the initial email requesting survey participation, as well as at the end of survey administration, a series of chi-square tests were conducted. As seen in Table 6, only one significant finding emerged: at the conclusion of the survey, the response rate for those offered the $100 incentive (16.2%) was significantly greater than the response rate for the control group (13.9%), χ^2 (1) = 3.93, p = 0.047. This finding may imply that the relationship between incentive amount and survey response is non-linear, as it is clear from the response rates and the statistical tests that "more" is not better: response rates did not increase as the amount of the incentive increased. However, given our large sample size combined with the marginal p-value, as well as the weak substantive impact of a 2.3% increase in response rate, the strength of support these data give for a non-linear effect may be suspect.

In addition to examining the effect of incentives on response rates, we also tested whether our experimental conditions had any effect on the quality of survey response. It is possible that an incentive might not change the probability that an individual will respond to a survey, but it might cause respondents to spend more time answering the survey. One way to test this hypothesis is to test item non-response between the experimental groups. If this is happening, we would expect lower item non-response for respondents in the incentive groups.

Our survey was adapted from the College Board's Admitted Student Questionnaire Plus and was comprised of six topics:

- Importance of college characteristics
- Characteristics ratings for the university
- Role of financial aid in the application process
- Images of the university;
- Number of applications mailed
- Demographic information

For each respondent, we calculated the number of survey items completed in each of the first

five sections separately and these scores served as dependent measures in a series of one-way ANOVAs. For the demographic variables, we recorded whether or not respondents supplied the requested information, and used the resultant binary data (0= did not supply; 1= supplied) in a series of chi-square tests.

In the series of one-way ANOVAs, we examined if (1) the number of survey items completed, or (2) the mean responses given varied across the survey conditions. As in earlier analyses, we tested the three main research questions outlined above.

Table 7 shows the results of the analyses conducted using the number of survey items completed or response/non-response as dependent measures. Significant effects of lottery incentives were only found for the importance of college characteristics and the number of college applications. We found no significant findings for the number of items completed for: the importance of characteristics, financial aid, or images of the institution, or the provision of demographic information.

With the exception of the $150 incentive, the mean number of items completed by each level of incentive was found to differ from the control group in the characteristics ratings section of the survey. Specifically, respondents in the $50, $100, and $200 incentive groups completed more items than respondents in the control group, with means of 16.8, 16.8, 16.9, and 16.3 (out of 17 items), respectively. Additionally, a significant effect of the overall impact of offering incentives was found for the number of characteristics ratings completed, with respondents offered incentives completing significantly more items (m = 16.75) than the control (m = 16.29). These findings suggest that the use of incentives may have caused respondents to complete a greater number of items specific to the university offering the reward, while more general survey questions were completed at a rate that was identical to respondents not offered an incentive. However, the substantive difference is quite small: about .5 items.

Analysis of the number of college applications found that respondents in the $200 incentive group applied to significantly more schools (m = 5.6) than the control (m = 5.0). This finding can be interpreted two ways. The first interpretation posits that the possibility of a large reward caused respondents in the largest incentive category to complete the survey more thoroughly. The second, and more plausible interpretation is that this finding is simply spurious - if the first interpretation were true, then we would have expected the analyses to reveal a larger number of significant findings.

To compare opinions of the survey groups on the importance of college characteristics in the application process and the characteristic rating section of the survey, we conducted a series of one-way ANOVAs using the mean response to each item as the dependent measure. Of the 204 tests conducted (34 survey items times 6 comparisons), only 6 (2.9%) were found to be significant. Because the number of significant effects was about what we might expect to find erroneously (at $p < 0.05$), we concluded that these (significant) findings were spurious.

In sum, the offer of a $100 gift certificate in a drawing increased the response rate by 2 percentage points, but there were no other significant differences between the control and

incentive groups. Given the very large sample size and p value, this is a weak finding, especially in terms of the substantive effect. In addition, it does not appear that offering larger amounts of incentives has a positive impact on response. There was some evidence that members of the incentive groups spent more time on the survey, as indicated by a slightly smaller item non-response rate. Again, the substantive differences were small.

Limitations

The chief limitation to our study is that the survey population is still not a college student population. It is possible that students' receptiveness to lottery incentives may change from their senior year in high school to when they enter college.

The overall response rate for the survey may also pose a problem. With such a low response rate, one could argue that interest in the survey was apparently so low that no incentive could have made a substantively large impact on response rates. Alternatively, the opposite argument could also be made. It is in low-interest surveys where incentives should make a difference and be effective, as respondents have few other reasons to participate. If respondents are very interested in a survey because of its content and thus are likely to respond, incentives may not have much of an impact beyond this interest.

It is also possible that given amount of the amount of spam and unwanted solicitation that are sent via email, we might have had more success with a paper survey. With a paper survey the lottery offer might have been more believable. We believe this had a minimal impact on our study, as our emails contained "institution.edu," sending a signal that we were members of a higher education institution (and thus increasing our credibility). The survey was also clearly located in the university domain.

The impact of the odds of winning a lottery is a further limiting factor for the study. In this study respondents could not estimate the odds of winning, as they had no idea how many other people had received an invitation to participate. For typical student survey respondents would have a rough idea of their odds of winning given the size of the student body and the number of prizes offered. Thus a lottery incentive might have more of an effect, as respondents would be better able to estimate the expected value of the incentive. Little, if any, research has been conducted on the impact of odds on lottery incentives and response rates, so this can only be a speculation as to what might have occurred if we had been able to conduct our analysis on a college student population.

Conclusion

Although the literature on incentives and response rates shows that post-payment of incentives in general and lotteries in particular have little or no impact on survey response, use of such lottery incentives appears common in institutional research. Our research is in line with previous research on the minimal effect of post-paid incentives. This raises a serious question of effectiveness and resource allocation. Given limited resources, should we be spending time and

money on awarding prizes, or on efforts proven to increase response rates, such as Dillman's (2000) method?

References

La Mar Adams, L. & Gale, D. (1982). Solving the quandary between questionnaire length and response rate in educational research. Research in Higher Education, 17(3), 231-240.

Boser, J. (1990). Surveying alumni by mail: effect of booklet/folder questionnaire format and style of type on response rate. Research in Higher Education, 31(2): 149-159.

Church, A. (1993). Estimating the effect of incentives on mail survey response rates: a meta-analysis. Public Opnion Quarterly, 57: 62-7.

Cote, L., Grinnell, R. and L. Tompkins (1986). Increasing response rates to mail surveys: the impact of adherence to Dillman-like procedures and techniques. Review of Higher Education, 9(2): 229-242.

Couper, M. (2000). Web surveys: A review of issues and approaches. Public Opinion Quarterly, 64: 464-494.

Dey, E. (1997). Working with low survey response rates: the efficacy of weighting adjustments. Research in Higher Education, 38(2): 215-227.

Dillman, D. (2000). Mail and Internet Surveys: The Tailored Design Method. New York: John Wiley & Sons.

Grosset, J. (1995). The biasing effects of bad addresses on information gathered by mail surveys. Journal of Applied Research in the Community College, (2)2: 179-191.

Groves, R., Cialdidn, R., and Couper, M. (1992). Understanding the decision to participate in a survey. Public Opinion Quarterly, 56: 475-495.

James, J., & Bolstein, R. (1992). Large monetary incentives and their effect on mail survey response rates. Public Opinion Quarterly, 56, 442-453.

Schiltz, M. (1988). Professional standards for survey research. Research in Higher Education, (28)1: 67-75.

Singer, E., Hoewyk, J. V., & Maher, M. P. (1998). Does the payment of incentives create expectation of effects. Public Opinion Quarterly, 62, 152-164.

Singer, E., van Hoewyk, J., & Maher, M. (2000). Experiments with incentives in telephone surveys. Public Opinion Quarterly, 64, 171-188.

Singer, E., Thurn, D. R. V., & Miller, E. R. (1999). Differential incentives: Beliefs about practices,

perceptions of equity, and effects on survey participation. Public Opinion Quarterly, 63, 251-260.

Smith, K. and T. Bers (1987). Improving alumni survey response rates: an experiment and cost-benefit analysis. Research in Higher Education, 27(3): 218-225.

Smith, T. (1995). Trends in nonresponse rates. International Journal of Public Opinion Research, 7,157-171.

Warriner, K., Goyder, J., Gjertsen, H., Hohner, P., & McSpurren, K. (1996). Charities, no; Lotteries, no; Cash, yes: Main effects and interactions in a Canadian incentives experiment. Public Opinion Quarterly, 60, 542-562.

Willimack, D., Schuman, H., Pennell, B., & Lepkowski, J. (1995). Effects of a prepaid nonmonetary incentive on response rates and response quality in a face-to-face survey. Public Opinion Quarterly, 59, 78-92.

Table 1. Number of Surveys Using a Lottery Incentive During a Typical Academic Year

Number of surveys with prizes	All schools		Private		Public	
	N	%	N	%	N	%
0	222	64%	53	50%	164	69%
1	75	22%	33	31%	41	17%
2	31	9%	15	14%	16	7%
3	11	3%	2	2%	9	4%
4	5	1%	1	1%	4	2%
5+	3	1%	1	1%	2	1%
Total	347	100%	105	100%	236	100%

Note: all schools contain respondents who did not identify their school?

Table 2. Types of Prizes Used

Prize	Total	Private	Public
Cash	22.1%	21.6%	22.9%
Gift certificate - national	11.5%	21.6%	2.9%
Gift certificate - local	23.0%	29.4%	17.1%
Gift certificate - school	57.4%	62.8%	54.3%
Travel prize	5.7%	11.8%	1.4%
Electronics	5.7%	7.8%	4.3%
Clothing	14.8%	13.7%	15.7%
Other type of prize	25.4%	17.7%	31.4%

Table 3. Perceived Impact of Lottery Incentives on Response Rates

	All respondents		Researchers not using incentives		Researchers using incentives	
	N	%	N	%	N	%
Greatly decrease response rates	0	0.0%	0	0.0%	0	0.0%
Slightly decrease response rates	1	0.5%	0	0.0%	1	0.8%
No effect at all on response rates	50	24.2%	39	46.4%	11	8.9%
Somewhat increase response rates	141	68.1%	41	48.8%	100	81.3%
Greatly increase response rates	15	7.3%	4	4.8%	10	8.9%
Total	207	100%	84	100%	122	100%

Table 4. Survey Experiment Groups

Group	N of ***High schools***	N of Students	Mean number of students per school
Control (no incentive)	706	1,983	2.9
Incentive - $50	706	1,712	2.5
Incentive - $100	706	1,960	2.8
Incentive - $150	706	1,784	2.6
Incentive - $200	705	1,866	2.7
Total	3,529	9,305	2.7

Table 5. Initial and Final Response Rates by Incentive Group

Group	After 1st email	After 3rd email
Control (no incentive)	4.6%	13.9%
Incentive - $50	5.4%	15.0%
Incentive - $100	5.3%	16.2%
Incentive - $150	6.0%	15.6%
Incentive - $200	5.8%	15.4%
All incentive groups	5.6%	15.6%
Total sample	5.4%	15.2%

Table 6. Hypotheses and Tests for Differences in Response Rates

	Dependent variable: response rate					
	After 1st email			After 3rd email		
Hypothesis	c^2	*df*	*p*<	c^2	*df*	*p*<
Control ? $50 ? $100 ? $150 ? $200	4.27	4	0.371	4.30	4	0.367
Control ? $50	1.26	1	0.262	0.97	1	0.324
Control ? $100	1.03	1	0.310	**3.93**	**1**	**0.047**
Control ? $150	3.45	1	0.063	2.15	1	0.142
Control ? $200	2.97	1	0.085	1.74	1	0.187
Control ? all incentive groups combined	3.17	1	0.075	3.36	1	0.067

Table 7. Hypotheses and Tests for Differences in Item Non-Response

Section	Number of items	Control ? $50 ? $100 ? $150 ? $200	Control ? $50	Control ? $100	Control ? $150	Control ? $200	Control ? all incentive groups combined
Importance of characteristics	17	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Characteristic ratings	17	**F(4,1397) = 3.96, *p* = .003**	**F(1,526) = 5.02, *p* = .026**	**F(1,584) = 7.37, *p* = .007**	**n.s.**	**F(1,555) = 8.49, *p* = .004**	**F(1,1400) = 9.97, *p* = .002**
Financial aid	2	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Images of institution	20	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Number of applications	12 max.	n.s.	n.s.	n.s.	n.s.	**F(1,555) = 4.05, *p* = .045**	n.s.
Gender	1	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Race/ethnicity	1	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
SAT	2	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
HS GPA	1	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

ADMINISTERING SURVEYS ON THE WEB: METHODOLOGICAL ISSUES

Heather S. Roscoe , Research Analyst
Dawn Geronimo Terkla, Executive Director
Gordon J. Hewitt, Assistant Director
Jennifer A. Dyer, Research Analyst

Office of Institutional Research
Tufts University

Introduction

The Internet is quickly becoming an integral part of everyday life. Since web-based surveys may be beneficial in terms of cost, time, and accuracy, more and more survey research (including that done in higher education and Institutional Research) is being conducted on the web (Underwood, Kim, & Matier, 2000). Thus, it is important to determine whether or not data obtained from web surveys are as valid and reliable as that of paper surveys, if response rates differ between paper and web surveys, and how ethical issues may effect the data collection, results, and reporting of the survey data. Moreover, ethical issues such as anonymity and/or confidentiality, the representativeness of the data, data analysis, and the use of quotations will need to be addressed when administering surveys on the web (Goree & Marszalek, 1995).

Only a handful of studies have explored how respondents react to web surveys, looking at issues such as bias in responses compared to paper and pencil surveys, or differences in overall response rates. With respect to response rates, Underwood et al. (2000) found that one mail (paper) survey had a notably higher response rate than another survey administered on the web. In addition, some research has indicated that females tend to respond in a greater proportion than males regardless of the method of survey administration (Underwood et al., 2000), while other results suggest that web surveys yield a greater proportion of male than female respondents (Tomsic, Hendel, & Matross, 2000). Moreover, recent findings suggested that certain ethnic minorities tend to respond at lower rates than do White, Asian American, and International Students regardless of the survey administration method (Underwood et al., 2000).

Other researchers have explored the issue of response bias or differences in responses based on method of survey administration. Baron and Siepmann (2000) found that responses to web surveys were very similar to the responses to paper surveys, but only if the formatting was exactly the same. Further, some studies have found that web surveys tended to receive significantly more favorable responses than did paper surveys (Carini, Hayek, Ouimet, & Kuh, 2001; Tomsic et al., 2000). Research conducted by Underwood et al. (2000) suggested that web surveys elicited less favorable responses than did paper surveys. Some findings have indicated that respondents tend to respond in a more honest fashion on the web (Turner, Ku,

Rogers, Lindberg, Pleck, & Sonenstein, 1998), while other findings suggest that respondents give more guarded/socially desirable responses on the web (Antons, Dilla, & Fultz, 1997).

The primary objective of this paper is to investigate if there are any psychological or methodological issues that would cause respondents to react differently to web surveys than they would to traditional paper and pencil surveys at a Doctoral/Research - Extensive university. In order to explore this area, data from two distinct projects will be used: 1) a survey of residential life and 2) an accepted applicant survey. An examination of the differences between the results obtained from paper and pencil surveys will be contrasted to those obtained via a web-based instrument. All of the previous studies used a cross-sectional design in their research.[1] The results were then compared across groups in order to detect differences. The research conducted for this paper departs from this methodology. This paper uses a longitudinal research design, where the trend of a given general population is analyzed at different data-collection points. This paper will examine how administering two annual surveys on the web impacted the response rates, as well as how web-derived responses compared to previous data obtained via paper surveys.

Experiment I-Residential Life

Method

For this study, data collected between 1997 and 2001 were used. Historically, residential life surveys were administered to all students living in staffed residential facilities. Data collection methods varied over time. In 1997, the survey was both administered and collected at a mandatory hall meeting. For the remaining years of paper administration (1998-2000), resident assistants (RAs) distributed the surveys to their residents and collected questionnaires when completed.

In 2001, the residential life survey was administered to all students (both undergraduate and graduate) living in all on-campus housing facilities. The survey instrument remained nearly identical each year. The survey contained many of the same questions with the same response options, as well as retained the same item order. In 2001 the page and print format changed to accommodate HTML, but the question format remained the same. The web survey was developed in-house using a commercially available software package and was published on a university server.

Emails with an embedded link to the survey were sent to all of the residential students in spring of 2001. This email briefly stated the purpose of the survey, the confidential nature of their responses, announced that students who responded would be entered in a raffle[3,] and

[1] That is, at a single time, paper (mail) surveys were sent to one group of respondents while another group of respondents were given a web-based survey.

[2]Three eligible students would be randomly selected would have $25 added to their meal cards. This money could be used at the campus bookstore, campus eateries, and campus convenience store.

contained a link to the survey. The email also stated that if the students had problems completing the survey on-line they could either pick up a paper copy at the Residential Life office, or contact the Office of Institutional Research (OIR) staff that developed the survey. A "reminder email" was sent to all students five days after the initial mailing. This reminder email thanked those that had already responded, and encouraged those who had not yet responded to do so. When the respondent clicked on the links contained within the emails, they were taken to a survey cover page. This cover page again ensured their confidentiality, the purpose of the survey, and the chances for winning a prize. In addition, RAs announced the survey during a regularly scheduled hall meeting, and hung flyers in the staffed housing facilities. The flyers contained the web address for the survey, and where the residents could obtain a paper copy of the survey.

Results

The web-based residential life survey experienced a notably lower response rate than past paper administrations. The respondents to the paper survey seemed to be an adequate cross-section of the university's population during those years, and the respondents of the web survey appeared to be fairly representative of this year's population. Approximately 48% of the items analyzed revealed that there were significant differences at the $\underline{p} = .05$ level between the responses of web respondents and those of paper respondents.

Demographics of Participants. The average response rate of paper respondents was 61.4% as compared to the web that yielded a response rate of 33.4%[4]. See Table 1. In order to allow students to report any difficulties responding to the survey, the email with the embedded link to the survey contained contact information for an OIR staff member. In addition, the email indicated that if they were unable to go on-line they could obtain a paper copy in the Residential Life Office. The project manager received only two emails from students unable to access the survey because they were using Macintosh computers. They were instructed to either use a PC to complete the survey on-line or to pick up a paper copy of the survey at the Residential Life Office. No paper submissions were received, and the social security numbers of those two students did not appear in the on-line submissions. There were 19 on-line submissions that did not provide a social security number. Therefore, it is unknown whether or not the two who initially encountered trouble with on-line submission did eventually submit a survey and merely did not provide a social security number. Due to the fact that this survey was not required, if students encountered problems filling out the survey, it is quite likely that they chose not to complete the survey and/or delete the email with the embedded link.

[4] It is quite likely that the software that was used to develop the survey instrument and to collect the responses may explain part of the reduction in response rate. The software would sometimes crash or freeze a respondents' computer. The program was only compatible with a limited number of browser editions, and the program was not compatible with Macintosh computers. We were able to obtain some limited anecdotal information.

Table 1

Residential Life Survey: Number of Respondents, Number of Potential Respondents, and Corresponding Response Rate

	Number of Respondents	Number of Potential Respondents	Response Rate
Paper	6,535	10,649	61.4%
Web	1,022	3,062	33.4%

The paper administrations yielded samples that averaged 43.9% male, and 56.1% female. See Table 2. In contrast, the sample of the web-based instrument was 41.1% male and 58.9% female. In order to determine if the method of survey administration affected the rate at which males and females respond, a χ^2 analysis was computed. Males seem to be slightly less likely to respond to the web version of the residential life survey than to the paper version. Conversely, females seem to be slightly more likely to respond to the web version than the paper version, a χ^2 analysis revealed that this trend was only approaching significance χ^2 (1, $\underline{N}$ = 6,945) = 2.631, $\underline{p}$ = .105. However, the undergraduate population has changed slightly. The average percentage of males during the paper administration years was 48.0%, and 46.6% for the web administration. Therefore, the decrease in percentage of males between paper and web surveys may be at least partially explained by a slight decrease in the male undergraduate population.

Table 2

Residential Life Survey: Number of Respondents by Gender

	Male		Female		Total	
	n	%	n	%	n	%
Paper	2,624	43.9%	3,353	56.1%	5,977	100.0%
Web	398	41.1%	570	58.9%	968	100.0%

Paper respondents averaged 67.3% White American, 15.9% Asian American, 8.6 % International Students, 4.2% Hispanic American, 3.8% African American, and 0.3% Native American. See Table 3. The distribution of students' race in the web sample differed somewhat: 64.1% White American, 18.2% Asian American, 6.4% Hispanic American, 5.7% African American, 5.4% International Students, and 0.2% Native American. To test for statistical differences, a χ^2 analysis was used. The χ^2 analysis revealed that there was a significant difference in the ethnic distributions between paper and web administrations χ^2 (5, $\underline{N}$ = 4,971) = 26.166, $\underline{p}$ = .000.

Table 3

Residential Life Survey: Number and Percent of Respondents by Ethnic Identity

	African American		Asian American		Hispanic American		Int'l		Native American		White American		Total	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Paper	155	3.8%	649	15.9%	171	4.2%	351	8.6%	12	0.3%	2,752	67.3%	4,090	100.0%
Web	50	5.7%	160	18.2%	56	6.4%	48	5.4%	2	0.2%	565	64.1%	881	100.0%

The web administration of the Residential Life survey elicited a higher percentage of African American, Asian American, and Hispanic American than did previous paper administrations. However, the ethnic distribution of the undergraduate population at the university has changed over the years. The African American population has grown from an average of 4.9% during paper administration years to 7.0% in 2001. The Hispanic American population has also grown from an average of 5.7% during paper administration to 7.8% this year. The Asian American population has decreased slightly from an average of 15.2% during paper administration to 14.9% during web administration. The ethnic distribution of the respondents to the paper survey looks as if it represents that of the entire undergraduate population from 1997 to 2000 fairly well, and the ethnic distribution of the web respondents seems to approximate that of the 2000-2001 undergraduate population at this university. Therefore, it appears that the change in the ethnic distribution at the university may account for the differences in the ethnic distributions between the paper and web surveys.

Response Pattern Analysis. In addition to differences in overall response rates, and the response rates by gender and ethnic identity, the pattern of responses to some of the individual items differed between the paper and web administrations of the residential life survey. An independent groups t test between web responses and paper responses was computed for questions that were asked on the 2001 survey, had been included on at least one paper survey from 1997 to 2000, and were based on the same scale. Analyses were conducted on a total of 29 survey items.

Eleven of the 29 total items (37.9%) received a higher mean rating from the web respondents than from the paper respondents, and five of those were significant. See Table 4. Fifteen items (51.7%) showed a decrease in mean response from paper to web, and nine of those were significant. Three items (10.3%) were given the same mean rating by both paper and web respondents, and none of these were significant. Overall, fourteen of the 29 items (48.3%) were found to be significantly different between paper and web responses. However, the differences in mean responses between paper and web administrations were small (typically less than 0.15 on a four point scale).

Table 4

Residential Life: Item Analysis of Response Means

	Increased from paper to web		Decreased from paper to web		No change from paper to web		Total	
	n of items	%	n of items	%	n of items	%	n of items	%
Significant	5	45.5%	9	60.0%	0	0.0%	14	48.3%
Not Significant	6	54.5%	6	40.0%	3	100.0%	15	51.7%
Total	11	37.9%	15	51.7%	3	10.3%	29	100.0%

Experiment II-Accepted Applicant

Method

Matriculating and non-matriculating surveys were administered to all students who were offered admission to the university each year and had a US home address. The surveys were not mandatory, no incentives for submission were offered, and no follow-ups such as a "reminder" postcard were used. These two survey instruments contained some of the same questions, and they each remained fairly consistent from year to year as far as questions asked, the order of items, and format.

Both surveys were shortened to facilitate on-line completion this year. Many of the remaining questions were the same as those from previous years. The surveys were developed using software available from a private web surveying company and was published on their server. Data from four years of paper administrations and data from the web administration 2001 were analyzed.

In 2001, emails containing a link to the surveys were sent to new students and non-enrolling students[5]. These emails briefly stated the purpose of the survey, but did not mention the estimate of the length of time required to complete the survey, nor were any incentives offered. Six days after the initial email, a "reminder email" was sent to all potential respondents. This reminder email thanked those who had already responded, encouraged those who had not already done so to respond, and gave contact information in case the potential respondent was having trouble loading the survey on-line.

When potential respondents clicked on the link embedded in the email, they viewed a cover page. This cover page briefly explained the purpose of the survey, the expected time to

[5] This year students with international mailing addresses were included in the pool of potential respondents due to the fact that both cost and complications with business reply envelopes would not be an issue

complete the survey (15 minutes for new students, 10 minutes for non-enrolling students), some web browser requirements, what to do if they were unable to load the survey or decided not to complete it, and were then thanked for their help. Only two emails from students who were experiencing difficulties with the web survey were received.

Results

The response rate to the new student survey web administration was somewhat lower than it had been to previous paper versions. Curiously, the non-enrolling student survey experienced a slight increase in response rate this year on the web than it had previously during paper administrations. Certain demographic features of the respondents such as gender and ethnicity were significantly different or nearly significantly different between the web and paper administrations, but can be partially explained by changes in the admitted student population and/or differences that were expected based on deliberate changes in the pool of respondents.

Thus, it seems as if both the paper and web administrations elicited responses from an adequate cross-section of the population at the time of administration. About 54% of the items on the new student survey that were analyzed revealed significant differences at the $p = .05$ level between the mean responses of paper and web respondents, while about 46% of the items on the non-enrolling student survey were significantly different at the $p = .05$ level between the two methods of survey administration.

Demographics of Participants. The average response rate for the new student paper survey was 67.3%, while the web administration in 2001 yielded a response rate of 57.5%. See Table 5. The paper versions of the non-enrolling student survey experienced an average response rate of 32.7%, and the web administration enjoyed a response rate of 33.6%. Thus, the new student survey's response rate fell nearly 10% with the web administration, while the non-enrolling student survey's response rate increased nearly 1% with the web administration.

Table 5

Accepted Applicant Survey: Number of Respondents, Number of Potential Respondents, and Corresponding Response Rate

	N of Respondents	N of Potential Respondents	Response Rate
New Student			
Paper	3,393	5,039	67.3%
Web	669	1,163	57.5%
Non-Enrolling Student			
Paper	3,569	10,915	32.7%
Web	677	2,015	33.6%

The paper administrations of the new student survey yielded samples that averaged 54.4% female, and 45.6% male, while the web administration of the survey yielded a sample that was 58.2% female and 41.8% male. See Table 6.

Table 6

Accepted Applicant Survey: Number of Respondents by Gender

	Male		Female		Total	
	n	%	n	%	n	%
			New Student			
Paper	1,533	45.6%	1,831	54.4%	3,364	100.0%
Web	248	41.8%	346	58.2%	594	100.0%
			Non-Enrolling Student			
Paper	1,450	40.9%	2,096	59.1%	3,546	100.0%
Web	231	40.0%	347	60.0%	578	100.0%

In addition, the paper administrations of the non-enrolling student survey yielded samples that averaged 59.1% female, and 40.9% male. The respondents to the web administration of the survey were 60.0% female and 40.0% male. A χ^2 analysis was used to test whether or not the method of survey administration affected the rates at which males and females responded.

This analysis revealed that for new students, females responded at a slightly higher rate to web surveys than to paper, and males responded at a slightly lower rate to web than to paper, but this trend was only approaching significance. χ^2 (1, $\underline{N}$ = 3,958) = 2.976, $\underline{p}$ = .084. For non-enrolling students, females responded at a slightly higher rate to web surveys than to paper, and males responded at a slightly lower rate to the web than to paper, and this trend was not significant χ^2 (1, $\underline{N}$ = 4,124) = 0.176, $\underline{p}$ = .674. However, the admitted students population had a slightly lower percentage of males this year than the average percentage of males in paper years. Therefore, the decrease in percentage of males between paper and web surveys may be at least partially explained by the slight decrease of males in the population.

The respondents to new student paper surveys averaged 69.0% White American, 16.1% Asian American, 8.0% Hispanic American, 6.8% African American, and 0.1% Native

American. The respondents to the new student web survey were 62.3% White American, 16.2% Asian American, 12.5% Hispanic American, 8.4% African American, and 0.6% Native American. See Table 7. Non-enrolling student paper surveys averaged 62.0% White American, 20.8% Asian American, 9.8% Hispanic American, 7.2% African American, and 0.2% Native American, while the web survey yielded 58.2% White American, 19.7% Asian American, 13.1% Hispanic American, 8.4% African American, and 0.6% Native American. To test whether or not ethnic groups responded at different rates to paper than web surveys, two χ^2 analyses were used. The χ^2 analyses revealed a significant difference in the ethnic distributions between paper and web administrations of the new student survey χ^2 (4, $\underline{N}$ = 3,287) = 18.490, $\underline{p}$ = .001, and the χ^2 analysis of ethnic distributions in the non-enrolling student survey was nearly significant χ^2 (4, $\underline{N}$ = 3,375) = 8.421, $\underline{p}$ = .077.

Both the new student and non-enrolling student web administered surveys elicited a higher percentage of African Americans and Hispanic Americans than did previous paper administrations. The ethnic distributions of the respondents to the paper surveys seem to represent that of the population of admitted students from 1997 to 2000 fairly well, and the ethnic distribution of the web respondents seems to approximate that of the population of admitted students in 2001. Therefore, it seems that the change in the ethnic distribution of admitted students may account for the differences in the ethnic distributions between paper and web surveys rather than the method of administration.

Table 7

Accepted Applicant Survey: Number and Percent of Respondents by Ethnic Identity

	African American		Asian American		Hispanic American		International		Native American		White American		Total	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
						New Student								
Paper	186	6.8%	442	16.1%	221	8.0%		0.0%	4	0.1%	1,898	69.0%	2,751	100.0%
Web	45	8.4%	87	16.2%	67	12.5%		0.0%	3	0.6%	334	62.3%	536	100.0%
						Non-Enrolling Student								
Paper	207	7.2%	597	20.8%	282	9.8%		0.0%	7	0.2%	1,780	62.0%	2,873	100.0%
Web	42	8.4%	99	19.7%	66	13.1%		0.0%	3	0.6%	292	58.2%	502	100.0%

Response Pattern Analysis. One hundred and fourteen survey items from the new student survey were analyzed for potential differences between paper and web respondents. Eighty of the 114 items (70.2%) received a somewhat higher mean rating from web respondents than from paper respondents, and 52 of those were significant. See Table 8. Thirty of the 114 items (26.3%) received a slightly lower mean rating from web respondents than from paper respondents, and nine of those were significant. In addition, four items (3.5%) showed no change from paper to web and these were not significant. Overall, 53.5% of the items received significantly different mean ratings at the p = .05 level from paper respondents than they did from web respondents. However, the differences in means between paper and web were small (generally less than 0.15 on a 4- or 5-point scale).

Table 8

Accepted Applicant: Item Analysis of Response Means

	Increased from paper to web		Decreased from paper to web		Showed no change from paper to web		Total	
	n of items	%	n of items	%	n of items	%	n of items	%
				New Student				
Significant	52	65.0%	9	30.0%	0	0.0%	61	53.5%
Not Significant	28	35.0%	21	70.0%	4	100.0%	53	46.5%
Total	80	70.2%	30	26.3%	4	3.5%	114	100.0%
				Non-Enrolling Student				
Significant	17	54.8%	1	14.3%	0	0.0%	18	46.2%
Not Significant	14	45.2%	6	85.7%	1	100.0%	21	53.8%
Total	31	79.5%	7	17.9%	1	2.6%	39	100.0%

Thirty-nine items on the non-enrolling student survey were analyzed for differences between the two survey methods. Thirty-one of the total 39 items (79.5%) received a higher mean rating from the web respondents than by the paper respondents, and 17 of those were significant. See Table 8. Seven of the 39 items (17.9%) were given a lower mean rating by the web respondents than by the paper respondents, one of which was significant. One item received (2.6%) exactly the same rating from both the web and paper respondents, and thus was not significant. Overall, 46.2% of the items were significantly different between paper respondents and web respondents at the p = .05 level. Again, the differences in means between paper and web were small (generally less than 0.15 on a 4- or 5-point scale).

Discussion

There are some limitations to the findings of this study. First, the web survey software package used to administer the residential life survey may have contributed to the low response rate due to the web browser restrictions, freezing, crashing, and the inability of Macintosh computer users to access the survey. It is likely that potential respondents were deterred by the difficulties they experienced, and since this was not a required survey, they did not make the extra effort to contact OIR for assistance or to go to the residential life office to obtain a paper copy. This may have confounded the response rate results.

Second, the accepted applicant surveys were dramatically shortened when being adapted for the web. Only essential survey items were retained. The new student survey web response rate was slightly lower than the average of previous paper years, and the non-enrolling survey response rate actually increased with the administration on the web. It could be that it was the shortening of the survey and not the administration via the web that kept the response rates approximately where they were with paper surveys.

A third limitation is that there was a large discrepancy in the number of respondents between the paper and web groups. Since four years of paper data were used in comparison to only one year of web data, the number of respondents in the paper group was four to five times greater than the number of web respondents. This may have over-exaggerated the meaning of the differences in mean responses between paper and web. In retrospect, using an alpha level of .05 may not have been strict enough due to the discrepancies in the numbers, and a level of .01 or .001 may give a more accurate assessment of the extent to which the means were different.

Overall, the two methods of survey administration seemed to obtain samples with similar gender and ethnicity distributions, as well as were representative of the populations from which they were sampled. For the most part, the differences in mean responses between paper and web were very small and were not any more unusual then year-to-year variation. The significant differences came in expected areas, and can be explained by slight changes in the population over time, or perhaps may be due to discrepancies in the number of respondents between the two methods of administration. Thus, it seems as if data obtained via web administration are comparable with that of paper administrations. In addition, administration of surveys on the web allows for sampling of segments of the population that were previously unavailable to us, such as students in un-staffed housing and students with overseas mailing addresses. The fact that the samples obtained via web surveys appear to be as representative as those from paper administrations, new populations can be sampled with web surveys, the data seems comparable, web surveys can be administered at a much lower cost compared to paper, and the increased expediency seen with web surveys, all point to the assessment that the web seems to be a valid mode of survey administration.

References

Antons, C. M., Dilla, B. L., & Fultz, M. L. (1997). Assessing student attitudes: Computer versus pencil. Paper presented at the Annual Forum of the Association of Institutional Research (37th, Orlando, FL, May 18-21, 1997). (ERIC Document Reproduction Service No. ED 410 866).

Baron, J., & Siepmann, M. (2000). Techniques for creating and using web questionnaires in research and teaching. In M. H. Birnbaum (Ed.), Psychological experiments on the internet (pp. 235-265). New York, NY: Academic Press.

Carini, R., Hayek, J. C., Ouimet, J. A., & Kuh, G. D. (2001). Student responses on web vs. paper surveys: A test of mode effects. Paper presented at the Annual Forum of the Association of Institutional Research (41st, Long Beach, CA, June 3–6, 2001).

Goree, C. T., & Marszalek, J. F., III. (1995). Electronic surveys: Ethical issues for researchers. College Student Affairs Journal, 15(1), 75-79.

Tomsic, M. L., Hendel, D. D., & Matross, R. P. (2000). A world wide web response to student satisfaction surveys: Comparisons using paper and internet formats. Paper presented at the Annual Forum of the Association of Institutional Research (40th, Cincinnati, OH, May 21-24, 2000). (ERIC Document Reproduction Service No. ED 445 633).

Turner, C. F., Ku, L., Rogers, S. M., Lindberg, L. D., Pleck, J. H., Sonenstein, F. L. (1998). Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. Science, 280, 867-873.

Underwood, D., Kim, H., & Matier, M. (2000). To mail or to web: Comparisons of survey response rates and respondent characteristics. Paper presented at the Annual Forum of the Association of Institutional Research (40th, Cincinnati, OH, May 21-24, 2000) (ERIC Document Reproduction Service No. ED 446 513).

STUDENT OUTCOMES: THE IMPACT OF VARYING LIVING-LEARNING COMMUNITY MODELS

Martha L. A. Stassen
Director of Assessment, Office of Academic Planning and Assessment
University of Massachusetts Amherst

Introduction

There is growing evidence that learning communities (LCs) can have a significant positive effect on a range of student and faculty outcomes (Lindblad, 2000; Matthews, Smith, McGregor, Gebelnick, 1997; Pascarella & Terrenzini, 1991; Smith, 1991; Tinto, 1998; Tinto, 2000; Upcraft, Gardner, and Associates, 1989). In the wake of these findings, higher educational institutions have adopted a variety of interventions that they identify as learning communities. These learning community models can vary greatly--from complex models with integrated curricula and on-site programming to more diffuse opportunities for students to meet together informally to discuss academic matters. The learning communities literature acknowledges this range of formats and has identified five different models (listed from the least coordinated to the most connected): 1. Linked Courses (two courses independent of each other, but with common students), 2. Learning Clusters (linked by content), 3. Freshmen Interest Groups (linked by theme), 4. Federated Learning Communities (faculty as the linchpin), and 5. Coordinated Studies Programs (where all the students' course credits are associated with an integrated, theme based, interdisciplinary curriculum, designed through intensive faculty collaboration (MacGregor, Smith, Matthews, & Gabelnick, n.d.; Snider and Venable, 2000). The fact that a great variety of models are all referred to as learning communities, and the reality that they are likely implemented with varying degrees of success, raise a number of evaluation questions that require attention.

While a wide array of programming and curricular strategies are labeled "learning communities", the relative impact of these different models can vary greatly. In their review of the literature on learning communities, Lenning and Ebbers (1999) identify a few studies that have looked at the impact of different learning community models and conclude: "*Well-designed* learning communities *emphasizing collaborative learning* result in improved GPA, retention, and satisfaction for undergraduate students" (p. 51, emphasis added). As the quote suggests, it is not yet clear whether it is primarily the more intensive models that shape the positive results attributed to learning communities. Indeed, Lindblad (2000) notes that most of the learning community research has been done on the "higher end" learning community model, Coordinated Studies Programs. The reality, however, is that most campuses cannot support these more coordinated (and resource dependent) models and have instead developed more modest learning communities. For these reasons, it is important to continue to compare the outcomes of various models to determine the extent to which these more modest models can achieve benefits similar to those that are clearly possible for more elaborate models.

In addition to pursuing the relative effect of more humble learning community models, further investigation is needed on the success of the full range of learning community implementations. Some of the most positive and widely disseminated results on the impact of learning communities appear to emerge from studies that do not necessarily include a full sample of the learning communities on the campuses studied. For example, in the extensive work of Tinto, Love, and Russo (1994), the researchers describe their selection methodology as follows: "In each institution, we selected a sample of learning community classes that in view of the program staff *best captured the intent of their program*" (italics added) (p.3). The question emerges, would the results of this study have been different had the full range of learning communities been analyzed? Again, are the generally positive effects of learning communities being driven somewhat by those with the most attentive implementation – and not the full range of learning communities that actually exist on campus.

Finally, the role of student self-selection into learning communities remains an issue in understanding their impact. In many cases, not all students on campus are involved in learning communities and students are not randomly assigned to the ones that exist. In studies where this is the case, and controls have not been put into place, the positive findings may be the result of students' own academic preparation and determination. For example, it is possible that students who are most motivated to succeed take advantage of the learning community opportunities and, as a result, retention and academic performance rates for learning communities are better because of individual student selection, not the program components themselves. Some recent studies (e.g., Venable and Snider, 2000) are particularly attentive to these concerns. However, much more work is needed in pursuing these issues.

Lenning and Ebber's (1999) summary of the findings of the Tinto, et. al. study provide the qualifications needed for understanding current learning community results – and suggest where future research must lead.

> Although we do not have complete assurance that the different models [in this study] were implemented with equal effectiveness or that the student groups were comparable on all potentially relevant variables, the results suggest that well-done, more concentrated, longer-term approaches to learning communities that involve faculty as active, intentional participants are more effective than others. (pp. 53-54)

Therefore, while the theory of learning communities seems to be supported in studies to date, the actual value (or impact) of any particular learning community design on any specific campus may vary substantially from the general findings. In addition, results may be affected by factors that have not yet been adequately controlled in all studies. To answer the questions raised above, this paper studies the relative impact of three learning communities models – each with different missions and slightly different structures – while controlling on a number of potentially influential variables. The three models represent the full range (and implementation) of learning communities on our campus; therefore avoiding the selection bias found in other studies of the learning community effect.

Learning Communities at UMass Amherst

Our public Research One University provides an ideal environment in which to study some of the questions raised above. We have supported learning communities on our campus for over 25 years and each year we enroll about a third of first-year students in them. These learning communities include:

1. **Residential Academic Program (RAP)**: RAP has been a presence on campus for over twenty years. It serves as the model on which our more recent variations are based. RAP students live in a common residence hall and enroll in a common freshman-writing course. In addition, they choose from a range of general education courses, some of which are taught in the residence hall. These general education courses are often large lecture courses with small discussion sections, led by Teaching Assistants, which are reserved for RAP students. RAP is open to all first-year students on a first-come, first-serve basis. Each year there are over 700 first-year students enrolled in RAP, one-half of which are undeclared.

2. **Talent Advancement Program (TAP)**: TAP is a variation of RAP that was first implemented over 10 years ago. It is a selective learning community that invites students with specific majors to enroll in a learning community program designed by their major department. TAP enrolls over 300 students each year in these programs. TAP students take at least two courses together and participate in a freshman seminar designed to introduce them to the work of the faculty. Most of these TAPs have faculty coordinators who work closely with students in the program.

3. **Honors College Learning Community**: Starting in fall 1999, the campus added an additional learning community experience, specifically for students admitted into the University's new Honors College. In this model, students sign up for one of a variety of small thematic learning communities and co-enroll in two honors general education courses per semester of participation. For the most part, these courses are small and faculty taught.

All three of the learning community models studied fall in to the "Linked Course" (and least coordinated) cluster in the general categorization of learning community models (Snider and Venable, 2000). There are, however, important differences in implementation across the three models. The first of these differences is in the criteria used to admit students into the programs. Two of the learning communities (TAP and Honors) are selective in their admission process while one, RAP, is open to students on a first-come, first serve basis. In addition, one model (TAP) is reserved specifically for students in selected majors, while Honors and RAP enroll a mix of declared and undeclared students.

The structures of the program also vary. All three models draw the foundation of their design from the RAP model, which means that students live together in a common residence hall and take at least two classes together. TAP and Honors, however, offer additional tailored options for their students. Each TAP program has a faculty sponsor who, to varying degrees,

shepherds students through their TAP year. Many TAPs also have major-specific seminars where faculty from the major meet with students. In general, students in TAP and Honors enjoy more direct faculty involvement and Honors students are more likely to have more small classes. In addition, all these models include a residential component.

In the aggregate, participation in these learning communities appears to have a positive effect on one-year retention. Table One shows the retention rates for learning community participants and non-participants across four recent first-year student cohorts. The one-year retention rate for students in learning communities is consistently higher.

INSERT TABLE 1

Even after controlling on entering characteristics for the most recent two cohorts (to account for some of the potential enrollment biases alluded to in the introduction above), the positive effect of learning communities maintains.

INSERT TABLE 2

The results in Table Two show that, controlling on a variety of entering socio-demographic and academic preparation characteristics (high school GPA, Math and Verbal SAT, gender, race/ethnicity, special support program involvement, school/college affiliation, and residency status), enrollment in learning communities on our campus has a significant positive effect on one year retention. As the odds ratio indicates, across the two cohorts students in learning communities are 35% to 37% less likely to leave than similar students not enrolled in a learning community.[1]

These general retention results, however, may mask some important differences in the effect of the three different living-learning community models we support. When retention rates for the two most recent cohorts are broken down by learning community type, a more complete picture of the effect of learning communities on student persistence emerges (Table Three). (Note that the Honors Learning Community is only included in the 1999 Cohort analyses because it was not in existence before that year.)

INSERT TABLE 3

For the 1998 Cohort, the TAP program retention rate is much higher than that for the RAP program, although both are higher than the rate for non-learning community students. Similar patterns are found in the 1999 Cohort, where both Honors and TAP rates are higher than those for RAPs, although all three are still higher than those for students not in a learning community. These patterns are not necessarily surprising, however, given that the Honors and TAP programs are selective learning communities that enroll some of the most well prepared

[1] To interpret logistic regression results, look at the "Odds Ratio" [Exp (ß)]. Where 1.00 means there is an even chance of staying enrolled and deviations from 1.00 indicate the increased (or decreased) chance of leaving. For example, in for the 1999 Cohort in Table 2., students in a learning community are .368 (or 36.8%) less likely to leave after their first year (odds ratio of 1.00-.632).

students on campus. The rates for these two programs clearly inflate the learning community retention rates when the rates for all learning communities are aggregated together.

As these analyses suggest, when a campus supports more than one learning community model on campus, answering the question, "Do Learning Communities 'work' on our campus?" becomes a complex question. One cannot assume that all models have similar effects nor can one assume the results will be completely consistent from year to year.

To provide some more in-depth answers to the "Do Learning Communities 'work'?" question on our campus, we conducted a multi-lensed analysis of the impact of the learning communities that serve our students. The comparative study looks at students at specific junctures in their first year. These junctures reflect the "input", "experience", and "output" components that Astin (1993) highlights as critical to consider in any assessment of institutional effectiveness. This study makes it possible to identify the specific and unique contributions of these different LC models and can help inform the development of additional living-learning communities. The study also has broader significance because it makes it possible to evaluate the relative success of the three living-learning models within a common institutional context. This perspective is unusual because campuses often support only one type of living-learning model, which makes it impossible to assess different approaches within a similar context. Finally, this study looks at all of the learning communities on campus, not just a selective few that represent the best-implemented versions of the models.

Methodology and Data Sources

The comparison of the living-learning experiences focuses on the following questions:

Do non-learning community students and RAP, TAP, and Honors LC students differ in preparation at entrance? (Because program participation is voluntary, a consideration of these "inputs" is crucial to a complete understanding of the impact of the programs).

What effect does enrollment in one of these learning communities have on students' academic performance and one-year retention (after controlling on entering characteristics and preparation), as compared to students not in a learning community?

At the end of the first semester, do the experiences of these four groups differ in significant ways?

Our investigation has relied on three data sources, described below, to capture information about students at different points during their first year on campus.

1. Longitudinal student data base information to document students' entering preparation and track their academic performance and enrollment patterns over the course of the first year.

The ACE-CIRP survey administered during summer orientation. This survey provides information regarding students' expectations and goals for college at entrance. (While about

90% of all entering freshmen take the survey, only 60% each year provide the necessary student identifying information to match with the institutional student database.

An end-of-first-semester phone survey. The sample for this survey, which was conducted at the end of the first semester, included a random sample of first-year students in learning communities and a random sample of those not in a learning community. The survey focuses on students' first-semester academic and social experiences (particularly those related to academic and social integration) and was developed in consultation with individuals responsible for designing and implementing the RAP, TAP, and Honors LC programs on campus. The Student Affairs Research, Information, and Systems Office, which does polling on students throughout the academic year, administered it. The response rate for the sample of students in an LC was 59% (n=477), for those not in a LC the response rates was 62% (n=328).

Data for this paper are drawn from the three most recent cohorts of first-year students (fall 1998, fall 1999, and fall 2000). Using these data we have employed chi-square and ANOVA analyses to compare the two groups on specific variables. To study the impact of these programs on academic performance and one-year retention, linear and logistic regression techniques are used.

Results

Inputs: Differences at Entrance

Do learning community and non-learning community students differ in college preparation at entrance?

Participation in TAP and the Honors LC is selective. Students are invited to participate based primarily on their high school performance and SAT scores. On the other hand, students are able to sign-up for RAP on primarily a first-come, first-serve basis. Given these differences in recruitment strategies, one would expect differences in academic preparation across the three learning communities, which might also lead to differences between those in learning communities and those not enrolled in these communities.

Table Four provides demographic and college preparation information for students in the different learning communities and for those not in a learning community for both the 1998 and the 1999 cohort.

INSERT TABLE 4

As this table shows, there are a number of significant differences across the groups of students. Looking first at demographic characteristics, while there is some variability, in general TAP, and Honors have fewer African American/Black and Hispanic students than are present in the non-learning community population. In 1999, TAP has more white students and honors has more students who refused to report their race/ethnicity. In both cohorts, RAP enrolls more female students than are enrolled in any of the other learning community or non-learning

community categories and in 1999, RAP and Honors have fewer out of state students enrolled.

Turning next to measures of academic preparation, the results of the two years consistently show that high school GPA and SAT scores are higher for those students in TAP and Honors than for students in RAP or those not enrolled in an LC. This pattern continues when we use the ACE-CIRP data to look at mother's and father's education. Students in TAP and Honors are more likely to have parents who attended at least some college.

These results show some important differences across learning communities. RAP students, in most ways, look more like the students not enrolled in a learning community than do students in TAP or Honors LC. For the measures used in this study, TAP and Honors students are better prepared than RAP or non-LC students. Again, this is not surprising given the selective nature of enrollment in these two programs. These "input" variables suggest the importance of controlling on these entering characteristics when exploring the outcomes of these varied LC programs.

Outcomes: Differences in Academic Performance and One-Year Retention
What relationship does enrollment in the various learning communities have with students' academic performance and one-year retention as compared to non-LC enrollment?

First Semester Grade Point Average

A central mission of the learning communities is to provide students with a learning environment that helps support their academic success. As the mean Grade Point Averages (GPAs) in Table Five indicate, students in all three learning communities do substantially better, on average, in their first semester than students not in a learning community.

INSERT TABLE 5

Of course, as the analysis of entering characteristics shows (Table Four), there are differences in academic preparation among these three groups. To determine the extent to which these differences in GPA maintain after academic preparation is taken into consideration, multiple regression was used to explore the role of RAP, TAP, and Honors LC enrollment on first semester GPA after controlling on high school GPA, SAT's, enrollment in special academic support programs, and a set of demographic characteristics (see Table Six).

These results show that, in both Cohorts, RAP, TAP, and Honors each has a significant positive effect on first-semester GPA even after all these entering characteristics are taken into consideration.

INSERT TABLE 6

One-Year Retention

Another primary purpose of the learning communities is to facilitate improved student retention. As mentioned earlier, for both Cohorts, the one-year retention rates show that all three learning communities have higher retention rates than is true for non-learning community students. (See Table Three). Again, it is important to control for the significant differences in students at entrance. To do so, we used logistic regression to control on entering characteristics and determine the effect of these three learning communities on one-year retention. Logistic regression makes it possible to calculate the "odds" of learning community participants leaving the University after their first year.

As Table Seven indicates, even after controlling on entering characteristics, RAP, TAP and Honors (in 1999) all have significant and strong effects on one-year retention. For example, in the 1999 Cohort, RAP students were 34% less likely to leave after their first year than similar students not enrolled in an LC, TAP students were 33.3% less likely to leave, and Honors students were 60.4% less likely to leave. It is not completely clear why the TAP effect decreased so substantially from 1998 to 1999 (although still a significant positive effect). There were no major programmatic changes so the phenomenon may have more to do with the introduction of the Honors program and the fact that some students who would have previously enrolled in TAP now enroll in the Honors LC.

INSERT TABLE 7

While retention itself is an important outcome, a further clarification of retention can provide additional insight into the effects of learning community enrollment. Students don't all leave for the same reason, and one of the biggest distinguishing characteristics is required withdrawal (where the University dismisses the student because of severe academic difficulty) versus voluntary withdrawal (where the student makes the decision to leave the University). Because academic dismissal indicates poor academic performance, the differences in college preparation across the various learning communities may lead to different learning community effects when type of withdrawal is used as the dependent variable. Table Eight shows the percent of students in each learning community category that left for one of these two reasons.

While the withdrawal and dismissal rates for all three learning communities are lower than for non-learning community students, the strength of effect is not consistent across all communities. In both cohorts, RAP students' voluntary withdrawal rate is more similar to non-learning community students' than are the TAP and Honors withdrawal rates. When it comes to dismissal rates, there is more difference between RAP and non-learning community students but, again, the difference is not as dramatic as it is for TAP and Honors students.

INSERT TABLE 8

Of course, academic preparation at entrance can be a key factor in predicting voluntary withdrawal versus dismissal. Tables Nine and Ten show the effect of the learning community programs on these two types of attrition after controlling on these entering characteristics.

INSERT TABLE 9

With respect to voluntary withdrawal (Table Nine), the learning community effect shows different patterns across the two cohorts studied. In the 1998 Cohort, TAP has a significant effect on decreasing voluntary withdrawal with TAP students 64% less likely to voluntarily withdraw. The RAP effect is also positive, but does not reach statistical significance. In the 1999 Cohort, RAP and TAP appear to have similarly strong effects (students in these two programs have just under a 30 percent greater chance of not withdrawing). The TAP effect does not reach statistical significance, probably because of the small sample of students in this category (N=29 TAP students as compared to N=87 RAP students). Honors learning community enrollment has an even stronger effect on voluntarily withdrawal. In the program's first year of operation (Cohort 1999), honors students were 52% less likely to voluntarily withdraw.

INSERT TABLE 10

Table Ten shows the same analysis for those students who were required to leave because of academic dismissal. In the 1998 Cohort, both TAP and RAP participation had a positive effect on not being dismissed. TAP students were 59% less likely to be dismissed and RAP students were 41% less likely to be dismissed. The results are fairly similar in the 1999 Cohort although the strength of the TAP effect decreases (38% reduced risk) and the odds ratio for TAP does not reach statistical significance. In this cohort RAP has an even stronger effect, with students in this learning community 49% less likely to be dismissed. Because there were no Honors students who were dismissed, Honors is excluded from the analysis.

Outcomes: Differences in First-Semester Experience

At the end of the first semester, do the experiences of these four groups differ in significant ways?

The analyses to this point have shown robust learning community effects for two student outcomes: first-semester GPA and one-year retention (voluntary and required withdrawal as well as retention overall). These results clearly suggest that there is something about the learning community experience as presented in all three of the models that has a positive effect on the first-year student experience. In the third part of this study, we explore how the student experience in the first semester might differ by learning community status.

For this part of the study, we surveyed first-year students in the 2000 Cohort at the end of their first semester at the University. The survey developed for this aspect of the study focused on experiential outcomes that the learning community literature, as well as those involved with learning communities on our campus, has suggested are the positive effects of learning communities. These fall into five categories (see appendix A for full description of all variables used in these analyses):

1. General Social Adjustment and Integration (i.e., degree of institutional commitment, involvement in extra-curricular activities, engagement with diversity).

2. Academic Integration

   a. Peer interaction around academic work (e.g., positive academic-related friendships, amount of time doing homework with peers, participation in group projects, etc.).

   b. Faculty interaction outside the classroom (amount of contact with faculty outside of class to discuss academic performance, discuss career options, socialize informally, and discuss course topics outside of class).

   c. Positive academic behaviors (e.g., being well prepared for class, participate in class discussions, amount of time spent doing homework, etc.).

   d. Positive academic climate (positive experiences in the classroom, perception of faculty being concerned about students, experiencing intellectual stimulation, having opportunities to integrate ideas across disciplines, etc.).

Table Eleven shows the results of two comparisons. In the left hand column, the results for students in any of the three learning communities are compared to those not in a learning community. In the right hand column, these same variables are compared across the three learning communities separately to determine if there are significant differences between the models.

First Semester Experiences: Learning Community – Non-Learning Community Comparisons

Looking first at the simple comparison between learning community participation and non-participation, there are few differences between the learning community and non-learning community experience in the first category of interest: general and social experience. While students in learning communities report greater institutional commitment, there are no significant differences in exposure to diversity (in values or race/ethnicity) or in ease of getting involved.

There are many more significant differences on the items reflective of academic integration. Students in learning communities are significantly more likely to have contact with peers around academic work, engage in group projects, report positive academic behaviors, study more hours, perceive a positive learning environment, and have had course assignments that required the integration of ideas. The amount of faculty contract is not significantly different.

INSERT TABLE 11

Because of the ongoing concern that the differences across learning community status in students' entering college preparation and demographic characteristics might influence the comparative results, we tested the robustness of these differences after controlling on high school GPA, SATs, gender, race/ethnicity, and residency status using multiple regression. For all of the academic integration variables (including faculty interaction) learning community participation had a significant positive influence. It did not, however, on the one significant item in the general and social experience grouping, institutional commitment (see Table Twelve for

the standardized beta coefficients of these analyses). The consistency of the relationship between learning community status and this set of academic integration indicators provides additional support for the positive effect of learning community involvement.

INSERT TABLE 12

First-Semester Experiences: Comparisons Between Learning Community Models

A central focus of this paper is on the potential for variability in outcomes across the different learning community models. The second column in Table Eleven makes these comparisons regarding students' first semester experiences. Among the general and social experience items there are some significant differences across the three learning communities. TAP students report slightly higher institutional commitment than do RAP or Honors students. On the other hand Honors students report more exposure to students with different values than do the other two groups. While in the aggregate, learning communities show little differences, this analysis does suggest distinctions across the three models.

Differences also exist among the academic integration items. TAP students have higher means on both of the peer interaction items and RAP students have lower means for both of the academic behavior items. There are no statistically significant differences across the three groups in faculty contact and the two academic climate variables.

Because of the substantial differences in academic preparation and demographic characteristics across the three programs, we again tested the relationship between the three learning community models and these experiential items after controlling on the same set of characteristics used above using multiple regression. Table Twelve provides a summary of these analyses.

Controlling on entering preparation and student demographics does not change the conclusions one can draw from the comparison of means although it further illuminates the role of the varying programs. With respect to the general and social experience variables, TAP enrollment has a small positive relationship with institutional commitment and Honors enrollment has a small positive relationship with exposure to diversity.

The relationships are much stronger on the academic integration items. All three learning communities have a positive relationship with peer interaction, but the effect is particularly strong for TAP students. Similar patterns are found for engaging in group work, although in this case, Honors participation is not significant.

While the mean differences for faculty interaction were not significant, in a multiple regression analysis, RAP participation emerges with a significant positive relationship. When it comes to academic behaviors, all three learning community models have a positive relationship with students' academic behaviors and the amount of time students spend studying. The RAP relationship with positive academic behaviors is particularly strong and TAP has a stronger relationship with hours spent studying.

Finally, on the two academic climate variables, TAP and RAP have significant positive

effects on both. Honors has no significant relationship with perceptions of a positive learning environment, but is related to students having course work that required the integration of ideas.

Discussion

The analyses for this paper were focused on addressing some of the gaps in the current research on learning communities. These include: (1) limited evidence of the benefits of more moderate learning community models (linked courses) as opposed to the "high end" learning community models (e.g., federated learning communities and coordinated studies programs); (2) limited studies on the full range of learning community implementations (as opposed to a focus on learning communities that best meet the program's goals); and (3) little information on the possible differential outcomes of varying linked course models. To explore these issues, this study uses a variety of outcomes, longitudinal analyses, and includes the full range of learning communities available on the campus. In addition, this study pays attention to differences in students at entrance.

The series of analyses conducted here provide consistent evidence that even the more modest models (implemented with varying resources, admissions criteria, and faculty involvement) have positive effects on a range of student outcomes. Across three different sets of dependent variables (first semester GPA, one-year retention, and a variety of academic experience measures) all three learning communities under analysis here have quite consistent positive effects. The same results are not found for students' general and social experiences, suggesting these learning communities emphasize academic integration more so than social integration.

While some of the differences may be numerically small, the overall pattern of effects provide compelling evidence of the positive influence learning communities can have on the first year student experience. These results indicate that even those programs that are minimally coordinated and varyingly implemented provide important academic benefits to a full range of students.

This study also focused on exploring the possibility of differential effects on student outcomes across the three learning community models. All three of the learning community models studied fall into the "linked course" category in the Snider and Venable (2000) framework (where students live together and take some set of classes together) but vary in admissions criteria and the amount of coordination and faculty involvement within this general framework. The TAP and Honors programs are more selective (and these differences are reflected in the student preparation and demographic characteristics of the students in the three programs). In general, students in TAP and Honors also enjoy more direct faculty involvement in their learning community and Honors students are more likely to have more small classes. Students in these two programs may also have a greater sense of affiliation with either the major (TAP) or with an elite college on campus (Honors). Rap students are less likely to have that sense of affiliation.

Given these differences in the three models studied, one might expect to see more

dramatically positive effects for the two learning communities that are more selective, more focused, and have more faculty involvement and more structured curricular options available to students.

The findings of this study, however, don't support this assumption. Looking first at the academic performance outcomes, while there is some variation across the two years of academic performance data studied, RAP and TAP show very similar positive effects on first semester GPA and both voluntary and required withdrawal. The Honors effect, however, is substantially stronger than the other two. The first-semester experience outcomes show more mixed results. After controlling on entering characteristics, the standardized beta coefficients for RAP are lower than those for TAP and honors on working with peers, hours spent studying, and opportunities to integrate ideas (although all still significant). However, the coefficients for faculty contact, positive academic behavior, and positive learning environment are as strong (and in some cases stronger) than those for TAP and Honors.

The results for faculty contact and positive academic behaviors are particularly interesting. While the simple means on these two items show little difference across the three learning communities (and in the case of academic behaviors, RAP's mean is substantially lower), once the academic preparation differences of students at entrance are controlled, RAP appears to have a strong positive effect. This suggests an interaction effect among the variables of interest. While the reasons behind this require further investigation, it may have to do with the fact that, as a group, RAP students are less prepared academically than the academic "stars" in TAP and Honors. While preliminary, these results might provide an indication that RAP improves the chances of interaction with faculty and engaging in positive academic behaviors for those students who may not be among the most highly prepared for college. The particularly strong and consistent effect that RAP has on preventing academic dismissal after one year (Table Ten) provides further support for the particular benefit of this type of learning community for those students who may be more at risk.

What is clear in the consistent pattern of results throughout this paper is that a variety of more humble learning community models can have a number of positive effects on the first-year student experience. These positive effects are not limited to those models that are highly coordinated or have extensive faculty involvement. These learning communities also work for students of varied academic preparation.

Limitations

While this study did take into account a range of student characteristics that might influence the outcomes under analysis here, the analyses do not necessarily include all potentially relevant input measures (e.g., student motivation, interest in working with others, etc.). In addition, the relationship between the first-semester experiences studied here and student

performance and retention have not yet been analyzed. The experience data comes from the fall 2000 cohort while the first semester GPA and retention data are from the 1998 and 1999 cohorts. Finally, this study does not necessarily shed light on the learning community processes that actually lead to the outcomes explored here.

Policy Implications

These findings seem to suggest that limited learning communities are better for students than no learning communities at all. Living together and taking a couple of linked courses with little other programmatic support has effects similar to those for models with more faculty involvement and curricular commonality. Also, student preparation need not be a barrier to learning community success since it appears that these models work for a range of students. Even simple structures that facilitate student interaction around academic work can have a positive effect for students of all preparation levels.

Bibliography

Astin, A. W. (1993). Assessment for Excellence. Phoenix, AZ: The Oryx Press.

Lenning, O.T. and Ebbers, L.H. (1999). The Powerful Potential of Learning Communities: Improving Education for the Future. ASHE-ERIC Higher Education Report, 26 (6).The George Washington University, Graduate School of Education and Human Development.

Macgregor, J., Smith, B. L., Matthews, R.S., & Gebelnick, F. (n.d.). Learning Community Models. Washington Center for Improving the Quality of Undergraduate Education.

Matthews, R. S., Smith, B. L., MacGregor, J., & Gebelnick, F. (1997). Creating Learning Communities. In J. Graff, J. Ratcliff, and Associates (Eds.), The Handbook of the Undergraduate Curriculum, (pp. 457-475). San Francisco, CA.: Jossey Bass Publishers.

Pascarella, E. T. and Terenzini, P. T. (1991). How College Affects Students. San Francisco, CA.: Jossey Bass Publishers.

Smith, B. L. (1991). Taking Learning Seriously: The learning community model. Liberal Education, 77 (2), 42-48.

Tinto, V. (1998). Colleges as Communities: Taking research on student persistence seriously. The Review of Higher Education, 21(2), 167-177.

Tinto, V., Love, A. G., & Russo, P. (1994). Building Learning Communities for New College Students: A summary of research findings of the collaborative learning project. University Park, PA: National Center for Postsecondary Teaching, Learning, and Assessment.

Upcraft, M. L., Gardner, J. N., & Associates. (1989). The Freshman Year Experience. San Francisco, CA.: Jossey Bass Publishers.

Table One. One-Year Retention Rates by Learning Community (LC) Status: All First-Time First-Year Students

1996 Cohort		1997 Cohort		1998 Cohort		1999 Cohort		2000 Cohort	
In LC Program	**No Program**	**In LC Program**	**No Program**	**In LC Program**	**No Program**	**In LC Program**	**No Program**	**In LC Program**	**No Program**
83.0%	77.7%	84.6%	76.7%	85.4%	78.6%	87.8%	81.4%	88.6%	81.5%

Table Two. One Year Retention Logistic Regression for Entire Cohort: Learning Community Participation **

	1998 Cohort N=3726		1999 Cohort N=3947	
	EXP B (Odds Ratio)	*Significance Level*	*EXP B (Odds Ratio)*	*Significance Level*
High School GPA	0.428	.000	0.519	.000
Verbal SAT	1.004	.442	1.001	.044
Math SAT	0.989	.112	0.998	.005
Gender (1 = Female)	0.839	.061	0.980	.833
Special Program Status (1 = In a Special Program)	0.571	.005	0.712	.138
Residency Status (1 = Out of State Student)	1.123	.214	1.200	.071
Race/Ethnicity*				
African American/Cape Verdean	1.258	.337	1.075	.789
Hispanic	0.812	.403	0.994	.982
Asian	1.508	.016	1.157	.435
School/College Affiliation*				
Humanities/Social and Behavioral Scies	0.786	.112	0.873	.360
Natural Science/Mathematics and Engineering	0.921	.532	0.864	.259
Applied Majors	0.730	.009	0.661	.002
College of Arts and Sciences Premajors	0.793	.090	0.841	.206
TAP/RAP/Honors (1 = In a LC)*	0.649	.000	0.632	.000

**Odds Ratio values of <1.00 indicate decreased risk of leaving

*Reference Group For Race/Ethnicity = White/Non Reporting Race/Ethnicity Students

*Reference Group For School/College Affiliation = Undeclared Students

*Reference Group For Residential Academic Program = Students in No Program

Table Three. One-Year Retention Rates by Type of Learning Community

1998 Cohort			1999 Cohort			
% Retained by Learning Community Mode(N)			*% Retained by Learning Community Mode (N)*			
None	**TAP**	**RAP**	***None***	**TAP**	**RAP**	**Honors**
78.6%	92.5%	83.4%	81.4%	89.7%	85.4%	93.7%
(2168)	(298)	(641)	(2201)	(315)	(644)	(210)

Table Four. 1998 Cohort and 1999 Cohort Entering Characteristics

	1998 Cohort				1999 Cohort				
Data Source: *Institutional Student Database*	**Sign.**	**None**	**RAP**	**TAP**	**Sign.**	**None**	**RAP**	**TAP**	**Honors**
Race/Ethnicity	P=.000 (N)				P=.000 (N)				
African American/Black	(182)	5.0	4.6	2.8	(142)	4.2	3.3	0.8	0.0
Asian American	(303)	8.9	4.1	7.5	(298)	7.4	7.3	7.1	7.1
Hispanic	(157)	4.0	4.4	3.7	(144)	3.4	5.8	1.1	1.8
Native American	(13)	0.4	0.3	0.0	(21)	.6	0.5	0.0	0.0
White	(2773)	70.1	76.3	73.9	(3000)	74.0	71.9	81.0	71.0
Non-Resident Alien	(48)	1.7	0.1	0.3	(37)	1.1	0.0	0.0	2.7
Non Reporting	(391)	9.9	10.2	11.8	(410)	9.2	11.2	9.9	17.4
Gender	P=.000 (N)				P=.000 (N)				
Female	(2020)	50.3	60.0	49.7	(2179)	51.7	61.1	54.0	53.6
Male	(1847)	49.7	40.0	50.3	(1873)	48.3	38.9	46.0	46.4
Residency	P=.549 (N)				P=.000 (N)				
Out of State	(1135)	29.8	27.9	28.6	(978)	26.4	17.7	24.4	18.3
In-State	(2734)	70.2	72.1	71.4	(3074)	73.6	82.3	75.6	81.7
Mean HS GPA	P=.000	3.13	3.13	3.47	p=.000	3.20	3.18	3.53	3.89
Mean Math SAT	P=.000	563	554	619	P=.000	562	542	625	664
Mean Verbal SAT	P=.000	551	559	618	P=.000	553	547	622	653
Data Source: *ACE-CIRP Survey**									
Father's Education	P=.088 (N)				P=.021 (N)				
No-College	(637)	27.4	25.4	20.6	(545)	26.0	23.3	20.4	15.7
Some College	(1780)	72.6	74.6	79.4	(1694)	74.0	76.7	79.6	84.3
Mother's Education	P=.004 (N)				P=.009 (N)				
No-College	(734)	31.5	30.3	20.5	(637)	29.2	30.9	20.6	19.4
Some College	(1695)	68.5	69.7	79.5	(1625)	70.8	69.1	79.4	80.6

*These data do not include the entire first-year cohort. While the majority of students complete the ACE-CIRP survey, a much smaller proportion give permission for their information to be matched to University records.

Table Five. 1998 and 1999 Cohort First-Semester College GPA

1998 Cohort				1999 Cohort				
Mean GPA by Learning Community				*Mean GPA by Learning Community*				
Sign.	**None *N=2745***	**RAP N=771**	**TAP N=322**	**Sign.**	**None N=2687**	**RAP N=752**	**TAP N=351**	**Honors N=220**
p=.000	2.51	2.71	3.08	p=.000	2.61	2.84	3.16	3.41

Table Six. First Semester GPA Predictive Model

	1998 Cohort N=3726		1999 Cohort N=3948	
	Standardized B	*Signific*	*Standardized B*	*Signific.*
High School GPA	.393	.000	.372	.000
Verbal SAT	.051	.000	.046	.010
Math SAT	.083	.005	.118	.000
Gender (1 = Female)	.078	.000	.066	.000
Special Program Status (1 = In a Special Program)	.074	.000	.029	.124
Residency Status (1 = Out of State Student)	-.011	.439	.010	.478
Race/Ethnicity*				
African American/Cape Verdean	-.033	.068	-.009	.579
Hispanic	-.010	.537	.014	.372
Asian	-.075	.000	-.025	.120
School/College Affiliation*				
Humanities/Social and Behavioral Sciences	.058	.000	.034	.028
Natural Science/Mathematics and Engineering	-.073	.000	-.063	.000
Applied Majors	.007	.682	-.001	.929
College of Arts and Sciences Premajors	.037	.021	.035	.025
Residential Academic Program*				
TAP	.083	.000	.071	.000
RAP	.077	.000	.108	.000
Honors LC	-----	-----	.045	.004
	R-Square=.235		R-Square=.232	

Table Seven. One Year Retention Logistic Regression Model: Comparison of Learning Communities Models **

	1998 Cohort N=3726		1999 Cohort N=3947	
	EXP B (Odds Ratio)	*Signific. Level*	*EXP B (Odds Ratio)*	*Signific. Level*
High School GPA	.438	.000	.536	.000
Verbal SAT	1.006	.312	1.002	.033
Math SAT	.991	.175	.998	.008
Gender (1 = Female)	.834	.052	.979	.826
Special Program Status (1 = In a Special Program)	.584	.007	.715	.142
Residency Status (1 = Out of State Student)	1.126	.203	1.200	.079
Race/Ethnicity*				
African American/Cape Verdean	1.274	.311	1.091	.747
Hispanic	.828	.451	1.000	1.000
Asian	1.543	.011	1.157	.433
School/College Affiliation*				
Humanities/Social and Behavioral Sciences	.803	.148	.875	.368
Natural Science/Mathematics and Engineering	.972	.828	.859	.249
Applied Majors	.771	.031	.657	.002
College of Arts and Sciences Premajors	.841	.211	.831	.182
Residential Academic Program*				
TAP	.375	.000	.667	.035
RAP	.746	.008	.660	.001
Honors LC	-----	-----	.396	.003

**Odds Ratio values of <1.00 indicate decreased risk of leaving

*Reference Group For Race/Ethnicity = White/Non Reporting Race/Ethnicity Students

*Reference Group For Residential Academic Program = Students in No Program

Table Eight. Voluntary and Required Withdrawal by Learning Community

	1998 Cohort			1999 Cohort			
	Percent Retained by Learning Community (N)l			*Percent Retained by Learning Community (N)l*			
	None	**RAP**	**TAP**	***None***	**RAP**	**TAP**	**Honors**
Withdrawal Status							
Voluntary Withdrawal	12.8% *(354)*	12.2% *(95)*	5.0% *(16)*	13.1% *(357)*	11.5% *(87)*	8.2% *(29)*	6.3% *(14)*
Academic Dismissal	8.6% *(236)*	5.4% *(42)*	2.5% *(8)*	5.9% *(160)*	3.3% *(25)*	2.3% *(8)*	0.0% *(0)*

Table Nine. *Voluntary* Withdraw Logistic Regression Model: Comparison of Learning Community Models (1=Withdrawn)**

	1998 Cohort N=3454		1999 Cohort N=3757	
	EXP B (Odds Ratio)	*Signific. Level*	*EXP B (Odds Ratio*	*Signific.e Level*
High School GPA	.670	.001	.687	.003
Verbal SAT	1.013	.091	1.001	.143
Math SAT	.981	.025	.999	.083
Gender (1 = Female)	1.021	.857	1.216	.076
Special Program Status (1 = In a Special Prog.)	.674	.143	.549	.031
Residency Status (1 = Out of State Student)	1.361	.005	1.465	.001
Race/Ethnicity*				
African American/Cape Verdean	.944	.859	1.063	.854
Hispanic	.797	.456	1.084	.783
Asian	.885	.620	1.331	.169
School/College Affiliation*				
Humanities/Social and Behavioral Scies	.902	.555	.933	.677
Natural Science/Mathematics and Engineering	.801	.197	.801	.151
Applied Majors	.820	.173	.640	.004
College of Arts and Sciences Premajors	1.024	.879	.820	.211
Residential Academic Program*				
TAP	.364	.000	.706	.107
RAP	.837	.168	.736	.023
Honors LC	-----	-----	.481	.023

**Odds Ratio values of <1.00 indicate decreased risk of leaving

*Reference Group For Race/Ethnicity = White/Non Reporting Race/Ethnicity Students

*Reference Group For Residential Academic Program = Students in No Program

Table Ten. *Dismissed* Withdraw Logistic Regression Model: Comparison of Learning Community Models (1=Dismissed)**

	1998 Cohort N=3284		1999 Cohort N=3481	
	EXP B (Odds Ratio)	*Signific. Level*	*EXP B (Odds Ratio*	*Signific. Level*
High School GPA	.196	.000	.252	.000
Verbal SAT	.995	.624	1.002	.098
Math SAT	1.002	.868	.997	.004
Gender (1 = Female)	.564	.000	.554	.001
Special Program Status (1 = In a Special Program)	.454	.003	1.132	.739
Residency Status (1 = Out of State Student)	.735	.047	.581	.013
Race/Ethnicity*				
African American/Cape Verdean	1.732	.083	.965	.932
Hispanic	.826	.623	.810	.638
Asian	2.772	.000	.822	.588
School/College Affiliation*				
Humanities/Social and Behavioral Sciences	.636	.088	.692	.195
Natural Science/Mathematics and Engineering	1.33	.133	1.084	.718
Applied Majors	.706	.068	.746	.216
College of Arts and Sciences Premajors	.491	.006	.899	.666
Residential Academic Program*				
TAP	.412	.021	.616	.208
RAP	.588	.005	.506	.003
Honors LC	-----		----	

Odds Ratio values of <1.00 indicate decreased risk of leaving

Honors LC not included in model because there were no dismissed Honors students.

*Reference Group For Race/Ethnicity = White/Non Reporting Race/Ethnicity Students

*Reference Group For Residential Academic Program = Students in No Program

Table Eleven. First-Semester Experiences: Mean Responses by Learning Community Affiliation

	2000 Cohort					
	No LC (N=328)	In a LC (N=477)		TAP (N=123)	RAP (N=257)	Honors (N=97)
Survey Items	***Sign.***	***Mean***	**Mean**	***Sign.***	**Mean**	**Mean**	**Mean**
General and Social Experience							
Level of Institutional Commitment (scale)	P=.014	3.39	3.49	P=.040	3.54	3.47	3.50
Amount of Exposure to Racial/Ethnic Diversity	ns	3.64	3.51	ns	3.54	3.47	3.57
Amount of Exposure to Diversity in Values	ns	3.32	3.39	P=.000	3.40	3.25	3.76
Ease of Getting Involved	ns	3.21	3.17	ns	3.16	3.19	3.15
Academic Integration Indicators							
Peer Interactions							
Extent of Academic Work with Peers (scale)	P=.000	2.84	3.24	P=.000	3.61	3.12	3.10
Number of Times Worked on Group Projects	P=.000	2.09	2.43	P=.000	2.77	2.35	2.23
Amount of Faculty Contact	ns	1.59	1.91	ns	1.58	2.04	1.95
Academic Behaviors							
Positive Academic Behaviors	P=.000	3.41	3.61	P=.045	3.65	3.56	3.70
Number of Hours Spent Studying	P=.000	10.90	12.94	P=.053	13.98	12.15	13.74
Academic Climate							
Experienced Positive Learning Environment	P=.000	2.58	2.72	ns	2.73	2.72	2.68
Course Work Required Integration of Ideas	P=.000	2.90	3.21	ns	3.21	3.19	3.25

Table Twelve. Summary of Regression Results: Standardized Beta Coefficients for Learning Community Effect on General and Social Experience and Academic Integration Indicators Dependent Variables. [1.]

	2000 Cohort			
	LC Effect			
	In a LC (N=477)	**TAP** (N=123)	**RAP** (N=257)	**Honors** (N=97)
Dependant Variables	**Std. B**	**Std. B**	**Std. B**	**Std. B**
General and Social Experience				
Level of Institutional Commitment (scale)	.064	.086*	.046	.042
Amount of Exposure to Racial/Ethnic Diversity	-.042	-.025	-.050	.002
Amount of Exposure to Diversity in Values	-.027	-.003	-.063	.101*
Ease of Getting Involved	-.029	-.035	-.013	-.054
Academic Integration Indicators				
Peer Interactions				
Extent of Academic Work with Peers (scale)	.257***	.401***	.152***	.188***
Number of Times Worked on Group Projects	.136***	.215***	.094*	.046
Amount of Faculty Contact	.089*	.036	.092*	.084
Academic Behaviors				
Positive Academic Behaviors	.168***	.128***	.160***	.119***
Number of Hours Spent Studying	.138***	.170***	.095*	.129***
Academic Climate				
Experienced Positive Learning Environment	.122**	.115**	.111**	.070
Course Work Required Integration of Ideas	.128**	.143***	.091*	.133**

[1.] Coefficients show relationship between learning community status and each indicator after controlling on Math and Verbal SAT, High School GPA, Gender, and Race/Ethnicity.

Indicators of Statistical Significance: *p≤.05; **p≤.01; ***p≤.001.

STUDENT SATISFACTION: MEASURES AND MEASUREMENTS

Emily H. Thomas, Director of Planning and Institutional Research
Nora Galambos, Associate for Institutional Research
Office of Institutional Research
State University of New York at Stony Brook

In a results- and measurement-oriented environment, the policymakers who oversee higher education, the parents who pay for it, and the students who make college choices look for evidence of institutional quality to differentiate institutions and guide decision making. This evidence includes objective outcome measures: Do students learn new facts or skills? Do they graduate? Are they subsequently successful in further education or careers? But subjective measures also indicate institutional quality: Do students have a rich and rewarding college experience? Are they satisfied? Do they believe they have learned and grown?

If students are viewed as consumers of higher education their satisfaction is important to institutional success, both because effective institutions should produce satisfied customers and because satisfaction supports the recruitment of additional customers. Indeed Astin concludes that "it is difficult to argue that student satisfaction can be legitimately subordinated to any other education outcome" (Astin 1993, 273). But student satisfaction is a complex construct influenced by a variety of characteristics of students and institutions (Benjamin 1994). To better understand these influences and opportunities to increase student satisfaction, we analyzed alternative measures of students' general satisfaction using alternative measurement techniques: multiple regression and decision-tree analysis.

Most previous research has focused on the characteristics of students and institutions that influence satisfaction (Astin 1993), identified the campus services with which students are more and less satisfied (Astin *et al*. 1987), or examined how satisfaction is related to other outcomes such as academic achievement (Bean and Bradley 1986) and retention (Tinto 1975 and subsequent research; Hatcher *et al*. 1992). In contrast, this analysis examines how students' satisfaction with specific aspects of their college experience influences their overall satisfaction.

Data. Data for this analysis are drawn from a student opinion survey at a public research university in spring 2000 (ACT 2000). The survey was administered to students in a representative sample of undergraduate classes and either completed in class or distributed in class for completion prior to the next meeting. Enrollment in the sampled classes totaled 15% of undergraduate enrollment; 64% of the sampled students responded, yielding a sample of 1,783. The survey collects data on students and their opinions on a broad array of experiences. For example, there are 44 measures of satisfaction with characteristics of the campus climate and environment, such as "your sense of belonging on this campus," "out-of-class availability of your

instructors," and "racial harmony at this college. There are also 35 measures of campus services and facilities, such and "library facilities," and "college social activities."

Satisfaction measures. In addition to collecting data on this array of satisfaction elements, the student opinion survey includes four questions that indicate students' overall satisfaction with their college experience:

1. Indicate your level of satisfaction with this college in general.
2. If you could start college over would you choose to attend this college?
3. What is your overall impression of the quality of education at this college?
4. It is likely that I will transfer to another college before next fall.

It is important for the study of student satisfaction to understand whether survey items such as these measure the same thing and what they mean. Many surveys include only summary questions like the first two to assess student satisfaction whereas our survey offers the opportunity to determine the specific elements of campus life that contribute most to these general ratings. Satisfaction with the quality of education offers an important comparison and highlights alternative consumer outcomes. To the extent that a college's mission is to provide education, its focus should be on ensuring customer satisfaction with education. Focusing on broader outcomes, such as students' general satisfaction, reflects the broader goal of providing a rewarding and pleasing environment. The likelihood of transfer provides a further perspective, on the assumption that dissatisfied students will "vote with their feet."

Three of the four satisfaction measures are moderately highly correlated, while the likelihood of transfer bears relatively little relationship to satisfaction:

Table 1. Correlation of General Satisfaction Measures

	Would choose this college again	Satisfied with the quality of education	Likelihood of transfer*
Satisfied with this college in general	.588	.590	-.202
Would choose this college again		.532	-.277
Satisfied with the quality of education			-.190

* Correlation for freshmen, sophomores and juniors only.

To understand why students respond differently to these general satisfaction questions we need to know what specific experiences their answers reflect. Identifying these experiences can also help faculty and staff set priorities for improving student satisfaction. Multiple regression provides one means of identifying the most important influences. Decision-tree analysis offers a richer account. Together the two types of analysis show that different indicators of general

satisfaction are influenced by a different array of student experiences, and different experiences affect the satisfaction of different types of students.

Regression analysis. Multiple regression identifies a small number of specific elements that "explain" a large proportion of the variation in students' overall satisfaction. Table 2 summarizes the results of stepwise regressions including all variables significant at the .001 level, listing standardized beta coefficients to indicate the relative effect of the explanatory variables.

Table 2. Predictors of Students' General Satisfaction*

	Satisfied with the quality of education	Satisfied with this college in general	Would choose this college again
Academic Experience			
Academic experiences [in the classroom]**	.201	.157	.160
Quality of instruction	.218		
Intellectual growth+	.219		
Preparation for life-long learning+		.109	.138
Social Integration			
Sense of belonging on campus		.252	.215
Personal security/safety on campus		.131	
College social activities			.124
Racial and ethnic diversity of students		.130	
Campus Services and Facilities			
Classroom facilities		.126	
Library services		.077	
Access to computing services and facilities	.081		
Academic advising services	.089		
Attitude of staff (non-faculty) toward students	.095		
Pre-Enrollment Opinions++			
Accuracy of pre-enrollment information		.180	.129
First-, second-, third-choice college	.097		.188
Good faculty was reason for choosing this college		.078	
Career prep. was reason for choosing this college	.133		
PERCENT OF VARIATION EXPLAINED (R^2)	49%	58%	37%

* The numbers listed are standardized beta coefficients which show the relative effect of the independent variables by measuring the number of standard deviations each changes for every one standard deviation change in the dependent variable, controlling for the effects of the other variables. Except as noted all questions are five-point scales ranging from "very dissatisfied" to "very satisfied."

** This important satisfaction indicator reads "How often have you been satisfied with your academic experiences at this college?" It appears in a section of the survey headed "Respond to the following questions about your classroom experiences at this college."

+ Five-point scale ranging from "none" to "very large."

++ The rating item is first choice=4, second choice=3, third choice=2, or higher choice=1 . The reason items are three-point scales ranging from "not a reason" to "major reason."

Out of 172 survey measures of students' characteristics, satisfaction, and campus experiences and plans, just 17 appear in the multiple regression models, and the three general satisfaction measures have different predictors. The variables predicting satisfaction with the quality of education are especially distinct. None of the social integration measures appear as predictors of satisfaction with the quality of education, whereas "sense of belonging" is the most important predictor of both the more general satisfaction measures. In each of the latter, a measure related to students' pre-matriculation attitudes and experiences is the second most important predictor. "Likelihood of choosing this college again" is heavily influenced by whether it was the students' first, second, third or other choice. Satisfaction with "this college in general" is heavily influenced by students' recollection of the information they received before enrolling, perhaps an indicator of the extent to which their expectations were met.

The variables that predict likelihood of transfer are completely different from the three general satisfaction measures. The small number of students expecting to transfer impedes the development of a statistically acceptable model from the survey data, but it appears that the significant predictors of transfer relate to career goals and the absence of the specific academic programs students desire.

The different things students mean when then respond to alternative measures of overall satisfaction should be taken into account when these measures are used in outcomes assessment. General satisfaction is not the same as satisfaction with educational quality. Moreover, the importance of pre-enrollment attitudes indicates that student satisfaction reflects inputs as well as college outcomes. The specific satisfaction predictors identified in this analysis offer guidance to program development aimed at increasing student satisfaction though the generalizability of these results is unknown. There is, however, no indication in these data that increasing student satisfaction will improve objective outcomes such as retention.

Decision-tree analysis. The data-mining technique of decision-tree analysis offers an alternative means of identifying the specific elements of the college experience that affect satisfaction. Multiple regression identifies influential variables as those most associated with the general satisfaction measures. The variables selected as predictors of student satisfaction are those that "explain" variance in the dependent variable. In contrast, decision-tree analysis identifies the important elements of students' college experience as those that most differentiate satisfied and dissatisfied students, based on chi-squared analyses that quantify differences in the frequency distributions of categorical variables.

The statistical algorithm, technically called CHAID (chi-squared automatic interaction detector), develops a "tree" of variables to explain variation in the dependent variable. It groups cases to create categories for each independent variable that are homogeneous with respect to the dependent variable. The algorithm selects from these categories the independent variable that identifies the groups most different on the outcome variable, based on a chi-squared test (or an F test for continuous variables). That variable becomes the first node in a tree with a branch for each category that is significantly different relative to the outcome

variable. The process is repeated to find the predictor variable on each branch most significantly related to the outcome variable, until no significant predictors remain.

The specific technique used in this analysis is "exhaustive CHAID" using SPSS Answer Tree, which employs an enhanced algorithm for grouping cases. Exhaustive CHAID more thoroughly examines all of the possible splits for each predictor.

Satisfaction with the quality of education. The experience that most distinguishes students who are satisfied with educational quality from those who are not is the perception of intellectual growth. Table 3 shows the frequency distributions that make this variable the first node of the decision tree. The chi-square statistic of 418.46 for intellectual growth versus satisfaction was higher than for any other variable.

Table 3. Association of Satisfaction with the Quality of Education and Perceptions of Intellectual Growth

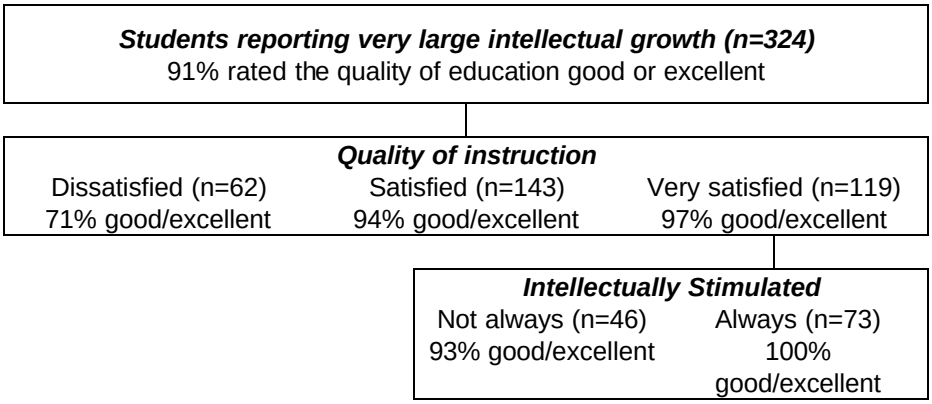
		Self-Reported Intellectual Growth (chi-square=418.46)			
	All students	Very large	Large	Moderate	Low/none
Percent rating the quality of education…					
Excellent	**.18**	**.46**	**.17**	**.06**	**.02**
Good	**.52**	**.45**	**.62**	**.50**	**.28**
Average	.23	.07	.18	.35	.35
Below average	.04	.02	.02	.06	.17
Very inadequate	.03	.00	.01	.03	.17
	1.00	1.00	1.00	1.00	1.00
Percent of sample	100%	19%	41%	33%	7%
n	1695	324	689	560	122

Overall, 70% of the survey respondents reported that the quality of education was excellent (18%) or good (52%), but the distribution of students' responses to the question about the quality of education was very different for students reporting high or low intellectual growth. For example, 19% of the students in the sample reported very large intellectual growth, and those students were very satisfied with the quality of education: 91% of them reported an excellent (46%) or good (45%) quality of education. In contrast, only 30% of the students reporting low or no intellectual growth rated the quality of education good or excellent. The decision-tree algorithm detected no significant difference between students reporting low growth and no growth and so grouped students in these two categories together.

Further branches of the decision tree identify different satisfaction predictors for different types of students, which is the greatest strength of the CHAID technique. It seems unlikely that very different kinds of students, for example those experiencing very high intellectual growth and those experiencing none, are influenced by the same aspects of the college experience. Pascarella and Terenzini (1979) highlighted students' heterogeneity when they investigated interactions between student characteristics and measures of social and academic integration, but this line of research was not pursued until recently when locating differences became of interest as a basis for market segmentation (Borden 1995). Previous research comparing the satisfaction of different types of students tested for differences among preselected categories, such as demographic groups (for example Sanders and Burton 1996), rather than searching for characteristics that distinguish satisfied and dissatisfied students. Both decision-tree analysis and stepwise regression as used in this study seek out significant patterns in a many-variable data set rather than testing specific hypotheses about the effect of preselected variables.

Table 4 identifies variables that influence the satisfaction of students experiencing very high intellectual growth by summarizing the frequency-distribution comparison in Table 3 for other independent variables. The percentage shown in each node is the percent of students in that group who rated the quality of education excellent or good. For example, among the 119 students reporting very large intellectual growth who were also very satisfied with the quality of instruction, 97% believed the quality of education was good or excellent, compared to only 71% of the 62 students dissatisfied with the quality of instruction. Moving further down the tree, 100% of the high-growth students very satisfied with the quality of instruction who were also consistently intellectually stimulated reported that the quality of education was either good or excellent. There were no significant subdivisions among the students less satisfied with the quality of instruction so these "limbs" do not branch further, given the restriction specified in the analysis that parent nodes have a minimum of 75 observations, and child nodes a minimum of 40 observations.

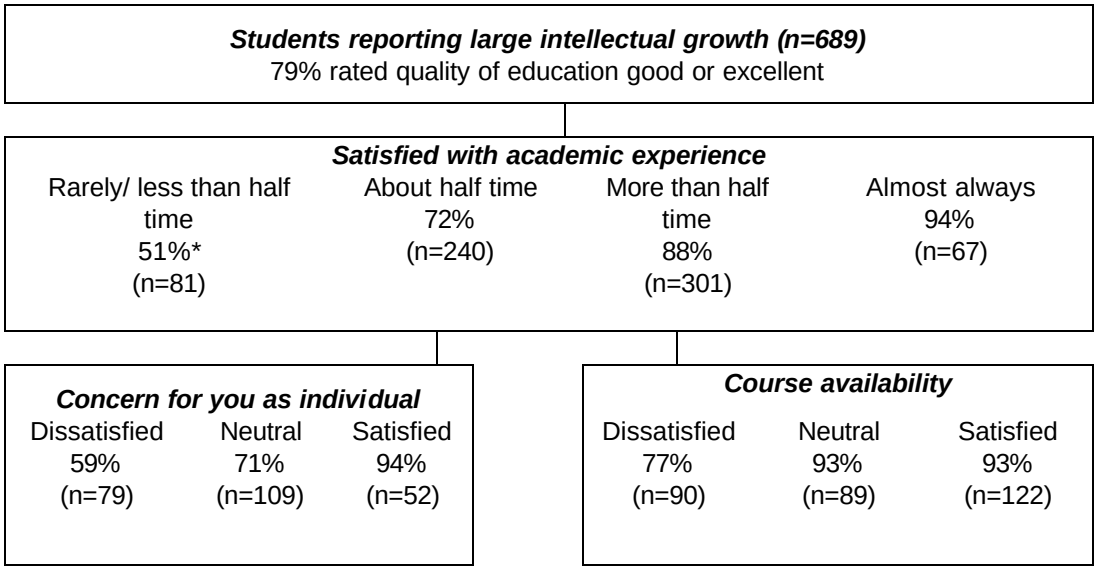
Table 4. Satisfaction with Quality of Education:
Students with Very High Intellectual Growth



These very high-growth students are high achievers. On a five-point scale ranging from 1=no academic growth to 5=very large academic growth, students with self-reported GPAs of 3.5 to 4.0 reported average intellectual growth of 3.79 compared to 3.38 for students with GPAs 2.0 to 2.5. They are the stars of the undergraduate population, the models of academic success, but on the diverse campus on which this study was performed they are just 19% of the student body. Analyzing and seeking to improve student satisfaction also need to consider the other 80% of the student population.

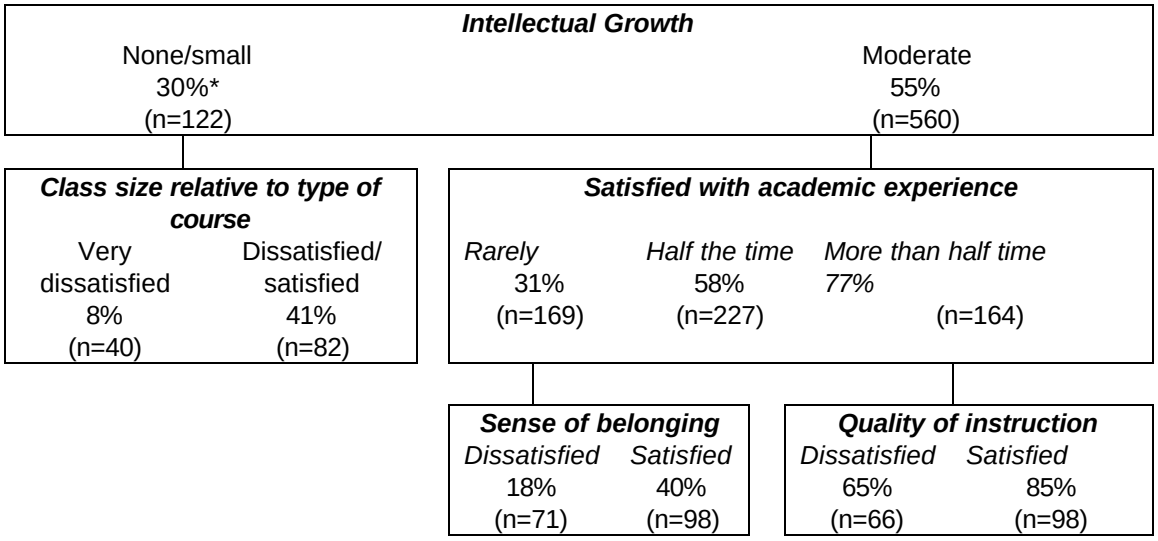
The rest of the decision-tree analysis is displayed in Table 5, which shows the tree branch for students who reported large (but not very large) intellectual growth (41% of the sample), and Table 6 which shows the students who reported low or moderate intellectual growth (40% of the sample). Not only academic factors, but also social and service variables are predictors of the satisfaction of these students with the quality of education. Perceptions that there is "concern for you as an individual" on the campus and satisfaction with course availability are associated with the relative satisfaction of students experiencing large intellectual growth. A sense of belonging and class size are important to students perceiving moderate growth or less.

Table 5. Satisfaction with the Quality of Education:
Students with Large Intellectual Growth



* Percent of students reporting that the quality of education was excellent or good.

Table 6. Satisfaction with Quality of Education:
Students with Low or Moderate Intellectual Growth

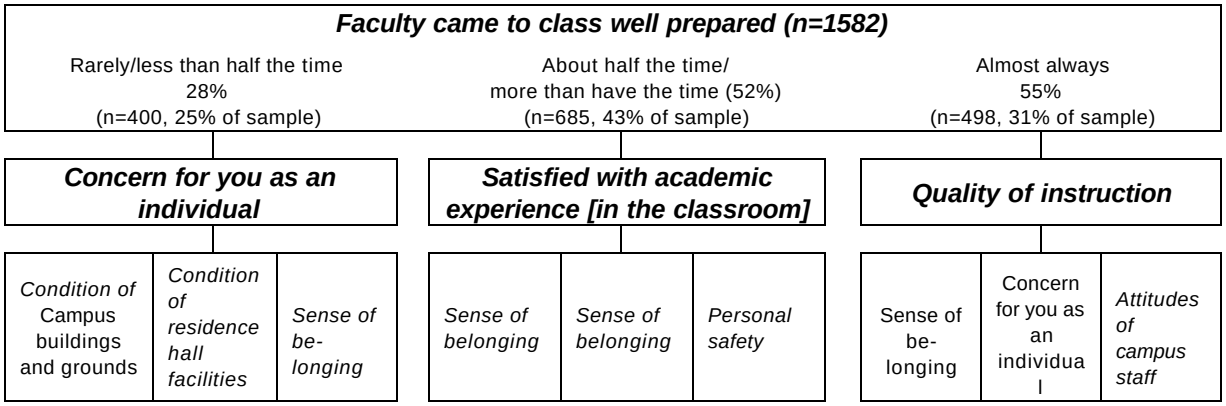


* Percent of students reporting that the quality of education was excellent or good.

Without displaying the tree structure, it is also worth noting that a different pattern distinguishes freshmen who are satisfied with the quality of their education, an important difference given the importance of students' first-year experience. Satisfaction with "preparation for a career" is the most important differentiating variable, perhaps reflecting the limited opportunity freshmen have had to grow intellectually or to recognize intellectual growth. More generally, self-reported intellectual growth increases with students class level.

Satisfaction with "this college in general." Decision-tree analysis of the factors that distinguish students more satisfied with "this college in general" also shows the importance of academic experiences to overall student satisfaction and supports the hypothesis that non-academic elements of the college experience are more important to less academically engaged students. As shown in the abbreviated decision-tree diagram in Table 7, it is how frequently students report having had "faculty who came to class well prepared" that most distinguishes their satisfaction. This is an ambiguous survey item whose interpretation requires investigation. Students' responses are likely to reflect differences in their ability to understand and absorb course material as much or more than the actual performance of faculty, since there is no reason to believe the students in different subgroups have been taught by different faculty. Nonetheless, the prominence of this variable indicates the importance to student satisfaction of classroom interaction with faculty. Faculty preparedness emerges as a more important determinant of student satisfaction than more general indicators such as satisfaction with "the quality of instruction" and "academic experiences."

Table 7. Satisfaction with This College in General



The second layer of the tree, identifying the variable that most distinguishes more- and less-satisfied students on each branch, displays the same relationship between student engagement and academic and non-academic experiences as the analysis of educational quality. The satisfaction of students who perceived the faculty as "almost always well prepared" were most differentiated by their perceptions of the quality of instruction. Those in the middle group, who perceived the faculty as generally well prepared, were further differentiated by satisfaction with their academic experience, an academic indicator but less specific than the quality of instruction.

The students least satisfied with the faculty are further distinguished by a non-academic variable, satisfaction with "concern for you as an individual." The nodes on this branch for students feeling there is little concern are the only ones on which satisfaction with campus facilities appears. These students may be so disengaged from the academic and social life of the campus that their satisfaction is influenced by perceptions of its physical characteristics.

Conclusion. Comparisons of students' experiences at a single campus provide several insights into student satisfaction. The generalizability of these conclusions needs to be tested, and comparison of students on different campuses could yield very different results. Nonetheless studying a single student body begins to identify aspects of the college experience that most affect student satisfaction. Within this population,

- Students' general satisfaction, satisfaction with the quality of education, and likelihood of returning to the same college measure satisfaction with different aspects of the college experience.
- Academic experiences most differentiate students who are more satisfied with college from those less satisfied, though a sense of belonging also contributes significantly to satisfaction. Satisfaction is heavily influenced by students' reaction to faculty in the classroom.
- Student satisfaction is significantly influenced by pre-college attitudes and as well as campus experiences. It reflects inputs as well as measuring college outcomes.

- It does not appear that satisfaction has an important effect on students' decision to transfer, though the data available from this study cannot offer strong evidence for or against this conclusion.
- The undergraduate population of a public research university is differentiated by varying perceptions of intellectual growth, and different types of experiences further influence the satisfaction of students with different experiences of growth.
- Academic diversity is a more important explanation of differences in student satisfaction than demographic diversity. No demographic variable emerged as a significant predictor.
- Differences in perceptions of campus facilities and services also appear to have relatively little effect on the varying satisfaction of students on a single campus.
- Non-academic aspects of college are more important to students who are less academically engaged than to those more engaged.
- Freshman satisfaction is most differentiated by perceptions of receiving career preparation, suggesting that faculty teaching first-year courses should help students understand how they represent a step toward students' career goals.
- The importance of programs that promote the social integration of freshmen is also suggested by these results, since intellectual growth increases with class standing and a sense of belonging is more important to the satisfaction of less intellectually-engaged students.

Faculty and administrators can examine student satisfaction in various ways. They can focus on improving the specific aspects of students' experience with which students are least satisfied. Or, if comparative data are available, they can address items on which the college is most different from its peers. Or, as in this project, they can identify those aspects of the college experience that most differentiate their own students' general satisfaction. Each of these approaches has limitations. How useful is it to know that the food service is rated low if all students dislike institutional food? How important is it to know that the campus bookstore is rated worse than others if that has little effect on students' overall college experience? And how important is it to know that differing perceptions of campus facilities does not affect satisfaction within the a particular student body, when all the students might be more satisfied in a more commodious setting? Despite these limitations, each approach can contribute to understanding student satisfaction by focusing attention on selected elements of the college experience that could be addressed in program development.

Intracampus comparative analysis also supports academic planning by profiling the student body as a non-homogenous population, differentiated not by demographics by intellectual experiences. Specifically, this analysis suggests the importance of developing programs and structures that integrate into campus life students who are relatively disengaged academically, since social integration appears relatively important to them. Honors colleges are undoubtedly exciting places for top students, but structures that promote social integration may have a greater effect on the satisfaction of less high-achieving students.

Regressions alone would not lead to these conclusions. Decision-tree analysis contributes a different perspective by identifying different predictive variables and differences within the student body. The validity and utility of this technique in studies such as this can only be proven by further use, but it appears worthy of further exploration as an alternative and complement to other statistical methods for drawing policy-relevant conclusions from many-variable surveys. Traditional analysis of differences in student satisfaction by demographic categories would yield little of interest from this survey, whereas the decision-tree analysis focuses attention on elements of students' experience worthy of campus discussion.

Bibliography

ACT, Inc. (2000). SUNY Student Opinion Survey. Iowa City, Iowa: ACT, Inc.

Astin, A.W. (1993). *What Matters in College? Four Critical Years Revisited*. San Francisco: Jossey-Bass.

Astin, Alexander, W Korn and K Green (1987). Retaining and Satisfying Students. *Educational Record* 68:1, 36-42.

Bean, J. P. and R. K. Bradley (1986). Untangling the Satisfaction-Performance Relationship for College Students. *Journal of Higher Education*, 57:4, 393-412.

Benjamin, M. (1994). The Quality of Student Life: Toward a Coherent Conceptualization. *Social Indicators Research* 31: 205-264.

Borden, V. M. H. (1995). Segmenting Student Markets with a Student Satisfaction and Priorities Survey. *Research in Higher Education*, 36:1. 73-88.

Hatcher L., K Kryter, J.S. Prus and V Fitzgerald (1992). Predicting College Student Satisfaction, Commitment, and Attrition from Investment Model Constructs. *Journal of Applied Social Psychology* 22:16, 1273-1296.

Pascarella, ET and P. T. Terenzini (1979). Interaction Effects in Spady and Tinto's Conceptual Models of College Attrition. *Sociology of Education*. 52: 4, 197-210.

Sanders, L and J D Burton (1996). From Retention to Satisfaction: New Outcomes for Assessing the Freshman Experience. *Research in Higher Education* 37:5, 555-567

Tinto V (1975). Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research* 45, 89-125.

LINKING LEARNING OUTCOMES TO STUDENT COURSE EVALUATION

Stephen W. Thorpe
Assistant Provost
Institutional Research, Planning, and Assessment
Drexel University

Abstract

Student ratings of course instruction are often criticized for collecting data of questionable use for assessment of student learning. This paper discusses the development of a local student course evaluation instrument that incorporates university-level learning outcomes.

Introduction

Student evaluation of course and instructor performance is a well-established practice in higher education. However, the use of student feedback is often criticized as being of questionable utility. Student assessment of instruction instruments (SAI) are often challenged as lacking validity and reliability. Moreover, while these instruments frequently ask students to evaluate the performance of the instructor, they seldom seek insights regarding the impact of the course on learning outcomes.

Student evaluation of instruction at Drexel University is a decentralized process that rests primarily at the college level; a standard university-wide evaluation form does not exist. The College of Engineering, for example, has developed an online course evaluation model that collects learning outcomes data from students enrolled in each course. In addition to standard questions asked for all courses, the instructor can add questions that ask students to evaluate specific learning outcomes for each course. The College of Business administers a standard evaluation instrument each term that collects data from students to inform faculty personnel decisions as well as to respond to their accreditation organization. Equally, the College of Information Science and Technology has administered their own college-level instrument for years for all courses within the college. Other academic departments and some individual faculty members have developed their own localized instruments as well.

As part of the self-study process for regional accreditation and assessment efforts on the campus, a specific set of learning outcomes were adopted at Drexel University for university-wide student outcomes assessment. The skills or abilities identified included: working effectively in teams; commitment to life-long learning; developing effective oral and written communication skills; identifying problems and developing viable solutions; and exhibiting an understanding of how solutions impact society. Each academic department was charged to develop departmental outcomes assessment plans that incorporated these learning objectives as appropriate in each course of study within the department.

To determine the extent to which skills and abilities are developed during the educational process, assessment offices typically rely on freshman, senior, and alumni surveys. While these types of surveys provide insights on the educational experience as a whole, they do not address skill and ability development at the course level. Therefore, a new course evaluation instrument was developed and piloted at Drexel that incorporates not only the typical questions regarding instructor performance, but also includes items that address university-wide learning outcomes.

Development of the Instrument Questions

The initial questions for the SAI instrument were adopted with permission from Dr. Thomas Angelo at the School for New Learning at DePaul University in Chicago, IL. Dr. Angelo is nationally recognized as a leading expert on course assessment strategies (see, for example, Angelo and Cross, 1993). The initial instrument, shown in Appendix A, consisted of 27 items categorized under the headings "About yourself," "About the course," "About the instructor," and "Summary questions". The university-wide learning outcomes were incorporated under the heading titled "About the course." These items included

- The course provided the opportunity to work in teams and team projects
- The course increased my desire to continue learning about this material
- The course provided the opportunity to learn/practice oral communication skills
- The course provided the opportunity to learn/practice written communication skills
- The course provided opportunities to identify problems and formulate solutions
- The course provided opportunities to develop viable solutions to problems
- The course allowed opportunities to exhibit an understanding of how solutions impact society, locally, nationally, and globally

A standard Likert-scale was used to collect responses for the questionnaire items. The "About yourself", "About the course", and "About the instructor" items were rated on a five-point Likert scale ranging from (1) never, (2) rarely, (3) sometimes, (4) usually, and (5) always. Items in the "Summary questions" were rated on a five-point Likert scale ranging from (1) extremely low, (2) low, (3) Adequate, (4) High, and (5) Extremely high.

The evaluation instrument also included open-ended questions to collect qualitative data for consideration by the instructor. These questions were

- In terms of your learning, what were the 3-4 best aspects of this course?
- What specific, practical changes can you recommend that might improve learning in this course?
- If a good friend asked whether you would recommend taking this course from the instructor, what would you recommend and why?

First Pilot Study, Winter 2001

The new SAI instrument was pilot tested in the winter, 2001 quarter in a sample of classes selected on a voluntary basis from three of the five colleges of study: Business, Design Arts, and Humanities. A total of 391 course evaluation surveys were completed in course ranging from anthropology, business law, management, psychology, and visual studies.

Pilot Study Courses
Winter, 2001

College/Department	Courses	Evaluations
Business		
Business Law	2	34
Management	1	59
Design Arts		
Visual Studies	16	176
Humanities		
Anthropology	1	20
Psychology	3	66
Sociology	2	36
Total	25	391

Factor analysis was used to determine whether students could differentiate among the different components of the course evaluation instrument, a practice that is commonplace when developing course evaluation instruments (Dolmans, D. et al, 1993; Gruetzemacher & Morris, 1992; Loftin, 1993; Ronco, 1999). Factor analysis is a useful statistical technique to identify a relatively small number of factors that can be used to represent relationships among sets of many interrelated variables (Norusis, 1985). For this study, a factor analysis utilizing principal components techniques was used. Factors with eigenvalues greater than one were selected for varimax rotation. The rotated factors are shown in Table 1.

Four factors were identified that explained about 59% of the variance. The first factor (explaining 33% of the variance) described the evaluation of instructor performance. The second factor described student preparation and work in the course, including items such as preparing for class, seeking help from the instructor when necessary, investing appropriate time in the course, and attending class sessions. Factor 3 loaded with all of the learning outcomes with the exception of commitment to life-long learning. The fourth factor described student interest in continued learning and interest in the course, essentially the dimension of commitment to life-long learning.

The factor analysis revealed Item 9 (the course promoted self-directed learning) did not load on any of the four identified factors. Accordingly, this item was deleted from the revised instrument.

In an interest to reduce the number of survey questions, a correlation matrix of the survey items was also reviewed to identify survey items with particularly high correlations (see Table 2). The correlation analysis revealed that two learning outcomes, "opportunities to identify problems/formulate solutions" and "opportunities to develop viable solutions to problems" were highly correlated (.77). These two items were combined for the next pilot study as "the course provided opportunities to identify problems and develop viable solutions."

In addition, items 7 and 8 that addressed course organization and clear communication of course objectives and requirements were highly correlated with item 20-"instructor communication of ideas and information" (correlations of .58 and .60 respectively). While item 7 was deleted from the instrument, the university is interested in knowing whether course objectives and requirements are clearly communicated. Accordingly, item 8 was retained on the SAI.

Finally, item 27, the "value" of what was learned, was highly correlated with item 26, the "amount" that was learned (correlation of .77), which suggested that students may be confusing the intended purpose of the two questions. Accordingly, item 27 was also deleted from the SAI.

Second Pilot Study, Spring 2001

The revised instrument (Appendix B) was pilot tested in similar classes that used the first instrument from the winter term at the conclusion of the spring, 2001 quarter. A total of 672 course evaluation surveys were completed in 35 courses.

Pilot Study Courses
Spring, 2001

College/Department	Courses	Evaluations
Business		
Business Law	2	59
Management	2	60
Design Arts		
Visual Studies	20	278
Humanities		
Anthropology	6	176
Psychology	3	76
Sociology	2	23

Total	35	672

Results of the factor analysis of the second pilot study were similar to the results from the first pilot study. Table 3 presents the factor loadings for the revised instrument and Table 4 shows the correlation matrix of the questionnaire items. Four factors were identified that explained 57% of the variance. The first factor again described the performance of the instructor, including communication of course objectives and requirements, providing useful feedback, communicating ideas clearly, and using class time effectively to promote learning. This factor also loaded with the items rating the effectiveness of the instructor and the amount learned in the course (items 22 and 23).

The second factor loaded with the learning outcomes. However, as with the first pilot study, the learning outcome of commitment to life-long learning did not factor with the other university student learning outcomes. Interestingly, this item again loaded on a fourth factor that might be considered student interest in the course material.

Future Plans

The results of the two pilot studies suggested that students could differentiate among the different components of the course evaluation instrument – evaluation of themselves, the instructor, and learning outcomes. Moreover, with the exception of measuring the commitment to lifelong learning, all of the learning outcomes loaded on one factor.

One of the assessment goals of the university is to collect data at the course level that measures the incorporation of the university-wide learning outcomes. By collecting this data each term, the university can describe where these learning outcomes are achieved during the educational experience at the course, department, college, and university levels. The pilot studies suggest that student learning outcomes can be effectively incorporated in the standard student evaluation of instruction instrument.

The revised instrument will be presented to the Faculty Senate during the 2001/02 academic year for consideration and adoption as a university-wide instrument for student assessment of instruction. Even if the instrument is not adopted in its entirety, the learning outcomes could be adopted for inclusion on instruments currently used at the department or college levels.

For example, the College of Business formed a committee two years ago to evaluate and develop a SAI instrument for the college. The selection of items for their evaluation instrument is in part governed by specific requirements of their national accrediting body.

Nevertheless, the committee has adopted the specific language of the learning outcomes items developed in the two pilot studies and will provide data on these items to the Provost's Office for university-wide reporting.

References

Angelo, T. and Cross, P. (1993). Classroom Assessment Techniques. New York: Josey-Bass.

Dolmans, D., Wolfhagen, I., and Schmidt, H. (1993). Validation of a rating scale for tutor evaluation in a problem-based medical curriculum. Paper presented at the Annual Meeting of the American Educational Research Association, Atlanta, GA. (ERIC Document Reproduction No. ED 362 523).

Gruetzemacher, R. and Morris, R. (1992). A factor analytic study of graduate education students' ratings of research methods instruction. Paper presented at the Mid-South Education Research Association, Knoxville, TN. (ERIC Document Reproduction No. ED 353326).

Loftin, L. (1993). Factor analysis of the IDEA Student Rating Instrument for introductory college science and mathematics courses. Paper presented at the Annual Meeting of the Mid-South Educational Research Association, New Orleans, LA. (ERIC Document Reproduction No. ED 366650).

Norusis, M. (1985). SPSSx Advanced Statistics Guide, Chicago, IL: SPSS.

Ronco, S. (1999). Desconstructing the student assessment of instruction instrument: Some psychometric issues. Paper presented at the AIR 1999 Annual Forum, Seattle WA. (ERIC Document Reproduction No. ED 433763).

Factor Loadings for Principal Components Analysis with Varimax Rotation
Winter, 2001 Pilot Study

Item No.	Questionnaire Item	Factor 1	Factor 2	Factor 3	Factor 4
1	I was well-prepared for each class session		0.751		
2	I asked the instructor for help/feedback when I needed it		0.659		
3	I invested enough time and energy to meet/exceed course requirements		0.769		
4	I participated actively and contributed thoughtfully during class sessions		0.697		
5	I attended class sessions (and related, required meetings)		0.595		
6	I practiced self-directed learning, including self-assessment and reflection		0.630		
7	The course was well-organized to help students learn	0.747			
8	The course objectives and requirements were clearly communicated	0.614			
9	The course promoted self-directed learning, including reflection/self-assessment	*	*	*	*
10	The course provided the opportunity to work in teams and team projects			0.697	
11	The course increased my desire to continue learning about this material				0.647
12	The course provided the opportunity to learn/practice oral comm skills			0.779	
13	The course provided the opportunity to learn/practice written comm skills			0.771	
14	The course provided opportunities to identify problems/formulate solutions			0.736	
15	The course provided opportunities to develop viable solutions to problems			0.712	
16	The course allowed oppt to exhibit an understanding of how solutions impact society			0.542	
17	The instructor provided me clear, useful, and timely feedback	0.82			
18	The instructor inspired interest/excitement in course material	0.785			
19	The instructor was available and helpful when asked	0.777			
20	The instructor communicated ideas and information clearly/effectively	0.813			
21	The instructor treated students, their ideas and opinions, with respect	0.617			
22	The instructor organized and used class time effectively to promote learning	0.779			
23	Before the course began, my level of interest in this course/topic was				0.704
24	Overall, I would rate the quality of my work in/for this course		0.560		
25	Overall, I would rate the effectiveness of the instructor	0.811			0.542
26	Overall, I would rate the amount I learned in this course	0.602			0.581
27	Overall, I would rate the value of what I learned in this course	0.528			
	Eigenvalue	8.83	2.81	2.62	1.62
	%Variance	32.70%	10.40%	9.70%	6.00%

Factor loadings of less than .40 absolute value are not displayed.
* No loadings greater than .40

Table 2
Correlations Among Questionnaire Items
Winter, 2001 Pilot Study

	Course Evaluation Item																
Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1.00																
2	0.43	1.00															
3	0.63	0.41	1.00														
4	0.49	0.53	0.52	1.00													
5	0.49	0.27	0.46	0.39	1.00												
6	0.42	0.48	0.42	0.39	0.24	1.00											
7	0.33	0.26	0.23	0.26	0.15	0.20	1.00										
8	0.28	0.14	0.22	0.15	0.16	0.12	0.70	1.00									
9	0.31	0.31	0.30	0.31	0.22	0.44	0.44	0.43	1.00								
10	0.02	0.12	0.08	0.17	0.13	0.18	-0.05	0.05	0.17	1.00							
11	0.25	0.25	0.24	0.21	0.15	0.25	0.51	0.46	0.45	0.14	1.00						
12	0.14	0.26	0.15	0.33	0.15	0.24	0.24	0.17	0.42	0.47	0.28	1.00					
13	0.06	0.01	0.06	0.12	0.05	0.08	0.19	0.17	0.24	0.46	0.23	0.55	1.00				
14	0.13	0.22	0.14	0.24	0.10	0.20	0.27	0.24	0.39	0.34	0.30	0.57	0.39	1.00			
15	0.17	0.26	0.21	0.24	0.15	0.26	0.34	0.33	0.45	0.33	0.36	0.54	0.42	0.77	1.00		
16	0.09	0.04	0.08	0.07	0.06	0.08	0.33	0.33	0.27	0.31	0.42	0.32	0.57	0.31	0.40	1.00	
17	0.33	0.34	0.25	0.28	0.19	0.19	0.60	0.53	0.39	0.02	0.46	0.24	0.22	0.29	0.35	0.30	1.00
18	0.25	0.24	0.19	0.21	0.12	0.18	0.64	0.53	0.36	0.04	0.50	0.31	0.27	0.27	0.33	0.40	0.69
19	0.29	0.37	0.23	0.26	0.20	0.21	0.56	0.44	0.40	0.04	0.38	0.24	0.17	0.30	0.35	0.20	0.72
20	0.25	0.26	0.18	0.22	0.17	0.11	0.58	0.60	0.39	0.08	0.40	0.29	0.30	0.27	0.34	0.39	0.67
21	0.26	0.32	0.27	0.24	0.25	0.24	0.42	0.41	0.34	0.15	0.29	0.19	0.15	0.22	0.28	0.21	0.58
22	0.21	0.25	0.21	0.26	0.10	0.16	0.54	0.43	0.29	-0.03	0.32	0.20	0.15	0.25	0.26	0.22	0.55
23	0.20	0.04	0.18	0.04	0.09	0.19	0.10	0.14	0.17	0.05	0.34	-0.03	0.00	0.00	0.12	0.16	0.03
24	0.38	0.27	0.43	0.29	0.22	0.29	0.21	0.23	0.32	0.14	0.30	0.21	0.18	0.20	0.27	0.23	0.22
25	0.19	0.29	0.14	0.24	0.07	0.15	0.60	0.47	0.35	0.03	0.44	0.30	0.19	0.30	0.33	0.33	0.63
26	0.26	0.23	0.24	0.21	0.05	0.18	0.67	0.57	0.38	-0.05	0.56	0.17	0.12	0.24	0.29	0.31	0.48
27	0.25	0.22	0.25	0.25	0.16	0.21	0.58	0.50	0.32	0.04	0.64	0.20	0.15	0.28	0.35	0.38	0.46

Table 2, continued
Correlations Among Questionnaire Items
Winter, 2001 Pilot Study

	Course Evaluation Item									
Item	18	19	20	21	22	23	24	25	26	27
18	1.00									
19	0.59	1.00								
20	0.67	0.63	1.00							
21	0.48	0.56	0.53	1.00						
22	0.56	0.51	0.55	0.40	1.00					
23	0.04	-0.04	0.02	0.02	-0.05	1.00				
24	0.26	0.19	0.20	0.18	0.15	0.26	1.00			
25	0.69	0.58	0.67	0.46	0.59	0.00	0.29	1.00		
26	0.52	0.44	0.47	0.31	0.44	0.20	0.39	0.55	1.00	
27	0.53	0.45	0.47	0.32	0.48	0.26	0.34	0.51	0.77	1.00

Table 3
Factor Loadings for Principal Components Analysis with Varimax Rotation
Spring, 2001 Pilot Study

Item No.	Questionnaire Item	Factor 1	Factor 2	Factor 3	Factor 4
1	I was well-prepared for each class session			0.577	
2	I asked the instructor for help/feedback when I needed it			0.674	
3	I invested enough time and energy to meet/exceed course requirements			0.673	
4	I participated actively and contributed thoughtfully during class sessions			0.622	
5	I attended class sessions (and related, required meetings)	*	*	*	*
6	I practiced self-directed learning, including self-assessment and reflection			0.595	
7	The course provided the opportunity to work in teams and team projects		0.696		
8	The course objectives and requirements were clearly communicated	0.718			
9	The course increased my desire to continue learning about this material	0.568			0.493
10	The course provided the opportunity to learn/practice oral comm skills		0.837		
11	The course provided the opportunity to learn/practice written comm skills		0.852		
12	The course provided opportunities to identify problems/develop solutions		0.759		
13	The course allowed oppt to exhibit an understanding of how solutions impact society		0.831		
14	The instructor provided me clear, useful, and timely feedback	0.741			
15	The instructor inspired interest/excitement in course material	0.82			
16	The instructor was available and helpful when asked	0.663			
17	The instructor communicated ideas and information clearly/effectively	0.866			
18	The instructor treated students, their ideas and opinions, with respect	0.707			
19	The instructor organized and used class time effectively to promote learning	0.777			
20	Before the course began, my level of interest in this course/topic was				0.484
21	Overall, I would rate the quality of my work in/for this course				0.628
22	Overall, I would rate the effectiveness of the instructor	0.785			
23	Overall, I would rate the amount I learned in this course	0.705			0.460
	Eigenvalue	6.48	3.45	2.03	1.25
	%Variance	28.10%	15.00%	8.80%	5.45%

Factor loadings of less than .40 absolute value are not displayed.
* No loadings greater than .40

Table 4
Correlations Among Questionnaire Items
Spring, 2001 Pilot Study

	Course Evaluation Item										
Item	1	2	3	4	5	6	7	8	9	10	11
1	1.00										
2	0.15	1.00									
3	0.44	0.33	1.00								
4	0.19	0.52	0.32	1.00							
5	0.26	0.10	0.25	0.16	1.00						
6	0.19	0.34	0.32	0.29	0.25	1.00					
7	0.00	0.10	0.02	0.21	0.08	0.00	1.00				
8	0.25	0.10	0.24	0.05	0.17	0.17	0.02	1.00			
9	0.16	0.15	0.23	0.19	0.11	0.24	0.02	0.47	1.00		
10	-0.01	0.28	0.10	0.34	0.05	0.13	0.54	0.11	0.19	1.00	
11	-0.02	0.16	0.04	0.20	0.05	0.03	0.50	0.05	0.12	0.66	1.00
12	-0.04	0.35	0.11	0.37	0.05	0.10	0.38	0.11	0.11	0.62	0.53
13	-0.01	0.14	0.07	0.19	0.06	0.04	0.44	0.13	0.19	0.60	0.70
14	0.17	0.28	0.22	0.19	0.11	0.20	0.03	0.46	0.36	0.16	0.12
15	0.21	0.22	0.23	0.19	0.25	0.25	-0.02	0.52	0.57	0.13	0.06
16	0.17	0.34	0.21	0.28	0.17	0.29	0.07	0.36	0.31	0.20	0.10
17	0.18	0.12	0.21	0.09	0.17	0.16	0.00	0.61	0.43	0.08	0.07
18	0.17	0.14	0.22	0.04	0.14	0.20	0.02	0.45	0.34	0.04	0.02
19	0.17	0.07	0.21	0.05	0.25	0.11	-0.07	0.52	0.40	0.04	0.04
20	0.10	0.03	0.09	0.06	0.11	0.18	-0.14	0.04	0.23	-0.05	-0.09
21	0.31	0.16	0.40	0.27	0.18	0.21	0.07	0.17	0.31	0.08	0.10
22	0.13	0.20	0.22	0.13	0.07	0.17	-0.06	0.53	0.55	0.12	0.08
23	0.16	0.13	0.26	0.15	0.14	0.21	-0.13	0.48	0.63	0.09	0.06

	Course Evaluation Item											
Item	12	13	14	15	16	17	18	19	20	21	22	23
12	1.00											
13	0.57	1.00										
14	0.19	0.15	1.00									
15	0.10	0.10	0.61	1.00								
16	0.21	0.09	0.59	0.55	1.00							
17	0.07	0.11	0.62	0.72	0.54	1.00						
18	-0.02	0.06	0.49	0.54	0.47	0.60	1.00					
19	0.08	0.10	0.51	0.63	0.48	0.72	0.52	1.00				
20	-0.04	0.01	0.11	0.13	0.05	0.08	0.05	-0.01	1.00			
21	0.08	0.11	0.19	0.23	0.18	0.20	0.20	0.19	0.16	1.00		
22	0.08	0.09	0.54	0.67	0.43	0.62	0.49	0.54	0.06	0.39	1.00	
23	0.10	0.13	0.45	0.59	0.37	0.53	0.35	0.48	0.18	0.42	0.71	1.00

Attachment A

DREXEL UNIVERSITY
Course Evaluation
WINTER 2001

The course evaluation is anonymous – do not put your name on either page of this evaluation. Please do the following:

- Use a #2 pencil
- Mark only one answer per item
- Fill in the ovals neatly and completely; do not make stray marks

About yourself
(A= Never, B= Rarely, C= Sometimes, D= Usually, E= Always

1. I was well-prepared for each class session
2. I asked the instructor for help/feedback when I needed it
3. I invested enough time and energy to meet/exceed course requirements
4. I participated actively and contributed thoughtfully during class sessions
5. I attended class sessions (and related, required meetings)
6. I practiced self-directed learning, including self-assessment and reflection

About the course
(A= Never, B= Rarely, C= Sometimes, D= Usually, E= Always)

7. The course was well-organized to help students learn
8. The course objectives and requirements were clearly communicated
9. The course promoted self-directed learning, including reflection/self-assessment
10. The course provided the opportunity to work in teams and team projects
11. The course increased my desire to continue learning about this material
12. The course provided the opportunity to learn/practice oral communication skills
13. The course provided the opportunity to learn/practice written communication skills
14. The course provided opportunities to identify problems and formulate solutions
15. The course provided opportunities to develop viable solutions to problems
16. The course allowed opportunities to exhibit an understanding of how solutions impact society, locally, nationally, and globally

About the instructor
(A= Never, B= Rarely, C= Sometimes, D= Usually, E= Always)

17. The instructor provided me clear, useful, and timely feedback
18. The instructor inspired interest/excitement in the course material
19. The instructor was available and helpful when asked
20. The instructor communicated ideas and information clearly and effectively
21. The instructor treated students, and their ideas and opinions, with respect
22. The instructor organized and used class time effectively to promote learning

Summary Questions
(A=Extremely low, B=Low, C=Adequate, D=High, E=Extremely high)

23. Before the course began, *my level of interest* in this course/topic was
24. Overall, I would rate the *quality of my work* in/for this course
25. Overall, I would rate the *effectiveness* of the *instructor*
26. Overall, I would rate the *amount I learned* in this course
27. Overall, I would rate the *value of what I learned* in this course

Attachment B

DREXEL UNIVERSITY
Course Evaluation
Spring 2001

Course: _____ Section: _____ Instructor: _____

Expected grade in this course: A B C D F

Is this course a requirement for your major? Yes No

About yourself
(A= Never, B= Rarely, C= Sometimes, D= Usually, E= Always, F=Not Applicable)

7. I was well-prepared for each class session
8. I asked the instructor for help/feedback when I needed it
9. I invested enough time and energy to meet/exceed course requirements
10. I participated actively and contributed thoughtfully during class sessions
11. I attended class sessions (and related, required meetings)
12. I practiced self-directed learning, including self-assessment and reflection

About the course
(A= Never, B= Rarely, C= Sometimes, D= Usually, E= Always, F=Not Applicable)

13. The course provided the opportunity to work in teams and team projects
14. The course objectives and requirements were clearly communicated
15. The course increased my desire to continue learning about this material
16. The course provided the opportunity to learn/practice oral communication skills
17. The course provided the opportunity to learn/practice written communication skills
18. The course provided opportunities to identify problems and develop viable solutions
19. The course allowed opportunities to exhibit an understanding of how solutions impact society, locally, nationally, and globally

About the instructor
(A= Never, B= Rarely, C= Sometimes, D= Usually, E= Always, F= Not Applicable)

20. The instructor provided me clear, useful, and timely feedback
21. The instructor inspired interest/excitement in the course material
22. The instructor was available and helpful when asked
23. The instructor communicated ideas and information clearly and effectively
24. The instructor treated students, and their ideas and opinions, with respect
25. The instructor organized and used class time effectively to promote learning

Summary Questions
(A=Extremely low, B=Low, C=Adequate, D=High, E=Extremely high)

26. Before the course began, *my level of interest* in this course/topic was
27. Overall, I would rate the *quality of my work* in/for this course
28. Overall, I would rate the *effectiveness* of the *instructor*
29. Overall, I would rate the *amount I learned* in this course

THE EFFECT OF FIRST-YEAR COLLEGE EXPERIENCES ON STUDENT PERSISTENCE: A CASE STUDY

Lillian Zhu
Director of Institutional Research
SUNY College at Brockport

Introduction

The purpose of this study is to examine the pre-entry attributes, the first-year academic performance, and institutional experience of the 1995 freshman cohort that graduated in no more than six years from a four-year public college located in a medium sized metropolitan city. Research has found that promoting student success in the first year is vital because approximately three-fourths of all dropouts leave during the first year (Tinto, 1993). During 1990s in this college (SUNY College at Buffalo where the author worked till recently), on average, more than 20% freshmen did not return after their first year of study. Meanwhile, cumulatively, only about 38% of the students graduated with a bachelor degree from the college by the end of sixth year. The high rate of first year attrition and the lower rate of graduation have direct impact on the cost productivity of the college and, therefore, puzzled the college administration for years.

The departure issue has been the object of empirical inquiry for decades. In recent years, the well known interactional theory of college students' departure postulated by Tinto (1975) has been questioned by scholars for the lack of empirical internal consistency (Braxton, 2000; Braxton, Sullivan, and Johnson, 1997) and the aggregated support or support by institutional type (Milem and Berger, 1997; Cabrera, Nora, and Castaneda, 1993). While researchers are reinvigorating Tinto's model through empirical affirmation, this study employed the logic of integration of the model to study the college persistence from three aspects, namely the pre-entry attributes, the first year academic performance, and the institutional experience.

The study attempted to answer the questions of great concern: (1) What are the factors to discern graduated and not-graduated groups? (2) Do these factors contribute to the college degree persistence at statistically significant levels?

Methodology

Data Source and Sample

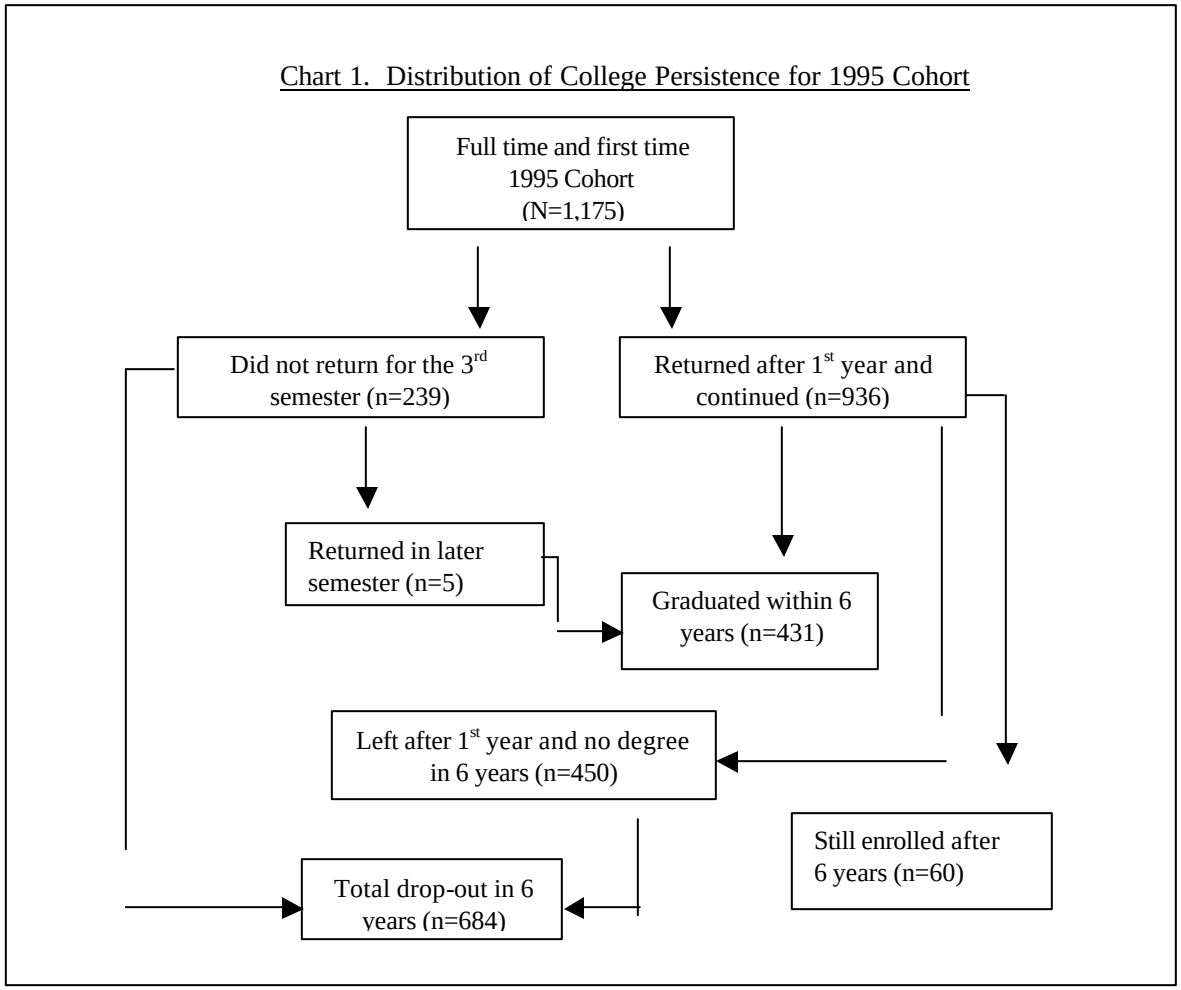
The study used 1995-2000 student data and 1998-2000 degree data of the college. Students who were first-time full-time freshmen and enrolled in fall 1995 were followed through the end of academic year 2000. A sample of 1,175 students formed the cohort for this study.

Measurement

Graduated/Not-Graduated. In this study, students who obtained their bachelor degrees from the college within a six-year period were identified as graduated, otherwise, not-graduated. Students who were temporarily out of the school after the first year but still managed to receive degrees belong to the graduated group (n=5). The status of neither still enrolled for the seventh year nor transferred to other schools is used as a criterion for the identification.

Variables for Pre-Entry and First Year Experience. The comparison between the graduated group and the not-graduated group was made on three sets of factors. The first set consists of the pre-entry attributes such as age, ethnicity, gender, higher school average, total SAT scores, and if entering college in the same year of high school graduation. The second set is about the first-year academic performance. It refers to number the registered credit hours, the number of remedial courses taken, the number of failed courses, and the first year GPA. The third set, or the institutional experience, includes the commitment to an academic major in the first year, participation of a special program (e.g. EOP), and the dormitory residency status. Unless specified, all the variables were used in their original format.

Chart 1. Distribution of College Persistence for 1995 Cohort



Results and Discussion

Of the 1,175 students in the cohort of the study, 431 (36.7%) obtained their bachelor degrees from the college within six years, while 744 (63.3%) did not. Chart 1 shows the distribution of the persistence. After the first year, 239 students did not return, an almost 20% first-year attrition rate. Five students left the college after the first year but managed to return and completed their degrees. Another five percent of the cohort (n=60) remained in the school after they started college six years ago.

Descriptive Analyses

The Chi-square analysis identified several factors that significantly differentiate the graduated from the not-graduated in each set of the variables (Table 1, 2, and 3). Table 1 indicates that females are more likely to complete their college degrees than their male counterparts. Those who were among the top 20% in high school GPAs are more likely to graduate from the college (75.9%). In addition, students who first entered college at an older age or did not attend college right after graduating from high school are less likely to persist.

Table 1. Graduated and Not-Graduated on Pre-Entry Attributes

Variable	Graduated (n=431)		Not-Graduated (n=744)		χ^2
	n	% within group	n	% within group	
Gender					7.574**
Female	278	64.5	419	56.3	
Male	153	35.5	325	43.7	
Age (1995)					14.735**
Less than 18	98	22.7	145	19.5	
18-19	315	73.1	546	73.4	
20 & up	18	4.2	53	7.1	
Ethnicity					0.064
Asian	11	2.6	26	3.5	
Black	52	12.1	137	18.4	
Hispanic	19	4.4	32	4.3	
White	322	74.7	495	66.5	
Others	27	6.3	54	7.3	
High School Avg.					18.994***
90+	37	8.6	37	5.7	
80-89	290	67.3	437	58.7	
70-79	70	16.2	172	23.1	
65-69	34	8.0	93	12.5	
Total SATs					2.735
1000 +	59	16.7	78	14.4	
800-899	172	48.6	261	48.2	
799 or Less	123	34.8	203	37.5	
HS Graduation Yr					12.256***
1995	399	92.6	638	85.8	
1994 or Earlier	32	7.4	106	14.3	

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

The factors of first year academic experience and performance also significantly differentiate the graduated from the not-graduated (Table 2). As reported in Table 2, where GPA was converted from numeric format into the ordinal letter grades, more than one third of those graduated took 14 or more credit hours (12 hours is the minimum to be a full time student)

Table 2. Graduated and Not-Graduated on First Year Academic Performance

Variable	Graduated (n=431)		Not-Graduated (n=744)		χ^2
	n	% within group	n	% within group	
1st Sem. Registered Hr					31.621***
Less than 14	271	62.9	578	77.7	
14 or More	160	37.1	166	22.3	
1st Yr Cum. GPA					54.855***
A	57	13.4	34	6.7	
B	236	55.4	209	41.0	
C	129	30.3	231	45.3	
D	4	0.9	36	7.1	
1st Yr Remedy Course					17.405***
None	336	78.3	399	67.3	
One	40	9.3	89	15.0	
Two or More	53	12.4	105	17.7	
1st Yr Failed Course					106.974***
None	295	68.8	259	43.7	
One	97	22.6	133	22.4	
Two or More	37	28.6	201	33.9	

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

compared to only one fifth for the not-graduated. Almost 69% of the graduated group received GPA at A or B while over 52% of the not-graduated group had C or D. The majority of the graduated students did not take remedial courses (78.3%) and did not fail any course (68.8%) during the first year, meanwhile a third of not-graduated students took at least one remedial course and/or had at least two failed courses. Therefore, a student who had better first year GPA, took fewer remedial courses, failed fewer courses, and earned more credit hours was more likely to graduate from the college.

The cohort experienced things that were associated with the institutional policies and the social economic background. Such experience might contribute to the college persistence (Table 3). More than one in every five students who enrolled in the EOP program ended up not graduated while 85% of those who entered the college through general admission graduated. No significant differences exist between graduated and not-graduated group in terms of major commitment or dormitory residency.

Hypothesis Testing

The rationale behind choosing the logistic analysis is the dichotomous nature of the dependent variable. In modeling, Pearson correlation analysis was performed to help select the quantitative components entering the regression model for hypothesis test. Instead of using first year GPA in letter grades, the numerical first year GPA was included in the logit model.

Table 3. Graduated and Not-Graduated on Institutional Experience

Variable	Graduated (n=431)		Not-Graduated (n=744)		χ^2
	n	% within group	n	% within group	
Admission Type					9.06*
General	370	85.8	590	79.3	
EOP	61	14.2	154	20.7	
Committed Major, 1st Yr					3.561
Yes	166	38.5	246	33.06	
No	265	61.5	498	66.9	
1st Sem. Live in Dorm					0.993
Yes	171	39.7	295	39.6	
No	260	60.3	449	60.4	

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

Each of the categorical independent variable was re-coded into dummies in order to capture the information contained in a categorization scheme. Then, this information was used in a standard regression estimation (Hardy, 1993). The number of dummies that can be created from a categorical variable must be the number of categories minus one. Omitting one category is to secure the mutual exclusiveness and exhaustiveness among the categories.

Three logistic models were applied to test if any factors from each set of the predictors, i.e. the pre-entry attributes, the first year academic performance, and the institutional experience, significantly contributed to the degree pursuing. Results of the analyses are reported in Table 4A. The parameter estimates (Table 4A) are the logit coefficients, which indicate the directions of the relationships between each pair of independent variable and the dependent variable respectively. The Wald χ^2 is a test statistic of the individual null hypotheses. The significance level specified with the Wald χ^2 tells whether an independent variable is significantly related to the dependent variable. The estimated odds ratio is the exponent of the parameter estimate, which interprets the magnitude of logit odds. The Model χ^2 is the global test statistic.

As reported in Table 4A, variable male is negatively related to the graduation status with statistical significance among the pre-entry attributes. This means that the males of 1995 cohort were less likely to graduate than their female counterparts. The odds ratio of male is 0.666, meaning that the odds for the male students to graduate is 66.6% as they are for the female students. The Model χ^2 is significant at 0.05 level.

First year cumulative GPA is positively and significantly related to the graduation status among the first year academic performance factors (Table 4A). The better the first year cumulative GPA is, the more likely the students would graduate. The odds ratio for the variable is 8.658, showing that the odds for students with higher first year cumulative GPA to graduate is more than eight times higher as they are for the students who had lower cumulative GPA in the first year. The Model 2 is significant at 0.001 level.

As for the institutional experience factors, enrolled in EOP program has a negative relationship with the graduation (Table 4A). The students enrolled in a special program, such as EOP, had lower probability of graduation. The odds ratio of 0.665 means the odds of

Table 4A. Logistic Analysis of First Year Experience on Graduation (N=1,175)

Variable	Parameter Estimate	Wald χ^2	Odds Ratio
Pre-Entery Attributes			
Intercept			
Age	-0.0380	0.4014	0.963
High School GPA	0.0022	0.3393	1.002
Total SATs	0.0007	3.1100	1.001
HS Grad. in Same Year	0.1204	0.1298	1.128
Male	-0.4066	8.0798**	0.666
Asian	0.3150	1.0893	1.370
Black	-0.2757	0.2688	0.759
Hispanic	0.0430	0.0134	1.044
White	0.5938	1.6689	1.811
	Model χ^2	17.712*	
	d.f.	9	
1st Year Acad. Performance			
Intercept			
Cumulative GPA	2.1585	109.0790***	8.658
# of Failed Course	0.2219	3.4542	1.248
# of Remedial Course	-0.3015	6.3028	0.740
	Model χ^2	707.505***	
	d.f.	3	
Institutional Experience			
Intercept			
Committed a major	0.0823	0.3567	1.086
Enrolled in EOP	-0.4084	5.0166*	0.665
Lived in Student Dorm	0.0066	0.0096	1.007
	Model χ^2	7.005	
	d.f.	3	

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

graduating for the EOP students is about two thirds those of general admission students. The Model χ^2 is not significant at 0.05 level.

In summary, the three logistic model tests have identified several factors that contributed to college persistence at statistically significant levels. Compared to the male students, females were more likely to persist. Students who were not admitted into special education programs were more likely to attain college degrees than those who enrolled in the EOP program. Moreover, the first year cumulative GPA contributed significantly to the college degree completion. The hypotheses of the study were thus partially accepted.

An additional logit model test was conducted in an attempt to investigate the aggregate effect of those factors on the college persistence. All of the independent variables were gathered and entered into one logistic model to test their effect on the dependent variable, the six-year graduation status. The results are reported in Table 4B. Four factors were found to significantly contribute to persistence, including the male gender (negative) and the first year GPA (positive). Hispanic students are 2.9 times as likely to persist compared to the non-Hispanics. For students with one more failed course during the first year, the odds of graduation are 72.6% those for those students with one less failed course. By and large, the results from the aggregate model support the findings from the early hypothesis tests.

Table 4B. Logistic Analysis of First Year Effects on Graduation (N=1,175)
(Aggregate Effect)

Variable	Parameter Estimate	Wald χ^2	Odds Ratio
Intercept			
Age	-0.1619	4.2821	0.851
High School GPA	-0.0221	0.2343	0.998
Total SATs	-4.99-E6	0.0001	1.000
HS Grad in Same Year	-0.1612	0.1668	0.851
Male	-0.3729	5.5478*	0.689
Asian	-0.1215	0.0419	0.886
Black	0.1375	0.1054	1.147
Hispanic	1.0630	3.6963*	2.895
White	0.2082	0.3652	1.232
Cumulative GPA	0.6642	18.5740***	1.943
# of Failed Course	-0.3198	9.5264***	0.726
# of Remedial Course	-0.0213	0.0131	0.979
Committed a major	-0.0520	0.1070	0.949
Enrolled in EOP	-0.4870	1.1974	0.614
Lived in Student Dormitory	-0.1210	2.1890	0.886
	Model χ^2	121.25***	
	d.f.	15	

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

The study went one step further to examine the group of students in the cohort who did not obtain their degrees at the end of the sixth year but still enrolled for their seventh college year (n=60). These students might have been out of the school for one or more semesters during their first six years. Table 5 summarizes the characteristics of this group and compares them with that of the graduated group. It shows that females accounted for 62% of still enrolled group and 82%

were white. About 13% of them were admitted into EOP in 1995. Among the first time full time cohort who entered the college in 1995, 87% graduated from high school in the same year. During their first college year, 80% did not take any remedial courses, but more than half of them failed at least one course. This group had higher average total SATs than that of the graduated group. All the still-enrolled had one thing in common, that is, for whatever the reason they took their own paces marching towards the graduation.

Table 5. A Brief Look of Students Not-Graduated but Still Enrolled (after six years in college)

Variable	Not-Graduated but Enrolled (n=60)		Graduated (n=431)	
	n	%	n	%
Gender				
Female	37	61.7	278	64.5
Male	23	38.3	153	35.5
Ethnicity				
White	49	81.7	322	74.7
Black	7	11.7	52	12.1
Others	4	6.8	57	11.2
Admission Type				
General	52	86.7	370	85.6
EOP	8	13.3	61	14.5
HS Graduation				
1995	52	86.6	399	92.6
1994 & earlier	8	13.4	32	7.4
1st Yr Remedial Taken				
None	43	79.6	336	78.3
1 course	8	14.8	40	9.3
2 & more courses	3	5.4	53	12.4
1st Yr Failed				
None	24	44.4	295	6.8
1 course	20	37.0	97	22.6
2 & more courses	10	18.6	37	8.6
Mean Scores	Mean	s.d.		
Age	23.9	1.0	18.5	2.5
High School Average	77.6	21.1	77.2	23.0
Total SATs	874.0	138.0	838.0	186
Cumulative Credit Hr	94.3	35.5	130.7	18.2
Cumulative GPA	2.7	0.5	2.8	0.4

Implications and Limitations

The findings of the study indicate that good academic performance in the first year positively

affects persistence. Taking fewer remedial courses implies a better pre-college preparation. Maintaining a GPA of 2.0 and above not only allows a student to claim an academic major and receive further financial assistance from EOP, but also assures the college path in front of his/her: you can do it! However, the GPA does not take into account of withdrawal and/or incomplete at the end of the first college year, one should exercise caution while interpreting the scores. As for college administrators, looking into the processing of GPA may be more important than reading GPA scores.

The study also found that female students were more likely to persist than male students in the college. The female students account for 59.3% who completed the college within six years. From the literature, the weight of evidence is clear because ability and socioeconomic status made women likely to be over-represented in the fields of education, social work and social sciences (for example, Jacobs, 1986; Polachek, 1978). Therefore, it is not surprising to find the higher persistence level for females since the college in this study offers more than half of its programs in education and is the first NCATE (National Council for Accreditation of Teacher Education) accredited institution in the state university system.

While the 1995 cohort study gave a broad view of the whole period of six college years, it focused on the first college year and the completion of bachelor degree only. As the study revealed, more than 58% of the 1995 cohort did not graduate and did not enroll any longer after six years, it is almost certain that more factors impact students' persistence and their departure after the first year in college. Many speculations have arisen that the mid-90's good job market and college's metropolitan surrounding pulled students away from the campus. But no statistics are available to support the claims. The data also lack information on the students who transferred to another school to pursue college degrees. To generalize the results, more work has to be done to explore the trend of the first year experience that impacts the college persistence.

The use of campus data often encounters the problem of data limitation. For this study, the model tests were limited to the availability of the data. Some of the variables that might have contributed to the college persistence according to the literature were not available, e.g. financial aid, student satisfaction, etc. Some data defects, such as missing information on the SAT scores and the dormitory residency variables, were noticed but are beyond the control of the IR office that maintains the campus data. Nevertheless, to IR practitioners, those data are still full of details, campus oriented, and suitable to serve various campus projects. However, use them with great caution.

The results of the study suggest that further studies may consider to discard the assumption of linear effect of predictors on the college persistency. Using non-linear model to prescribe the regression may produce a more realistic curve than applying linear perfection. Checking out the non-linear effects of predictors can be done through examining the specific odds level of the variables.

References

Baxton, J. M., L. A. Lien (2000). The viability of academic integration as a central construct in Tinto's interactionalist theory of college student departure. In J. M. Braxton (ed.), Reworking the Student departure Puzzle. Vanderbilt University Press: Nashville.

Baxton, J. M., A. S. Sullivan, and R. M. Johnson (1997). Appraising Tinto's theory of college student departure. In J. C. Smart (ed.), Higher Education: A Handbook of Theory and Research, vol. 12, pp. 107-164. New York: Agathon Press.

Cabrera, A. F., A. Nora, and M. B. Castaneda (1993). College Persistence: structural equations modeling test of an integrated model of student retention. Journal of Higher education, 64:123-139. Hardy, M. A. (1993). Regression with Dummy Variables. Sage Publications: Newbury Park.

Jacobs, J. (1986). The sex-aggregation of fields of study: Trends during the college years. Journal of Higher Education, 57, 134-154.

Milem, J. F., and J. B. Berger (1997). A modified model of student persistence: Exploring the relationship between Astin's theory of involvement and Tinto's theory of student departure. Journal of College Student Development, 38:387-400.

Polachek, S. (1978). Sex difference in college major. Industrial and Labor Relations Review, 31, 498-508.

Tinto, V. (1993). Leaving College: Rethinking the Cause and Cures of Student Attrition, 2nd ed. Chicago: University of Chicago Press.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. Review of Educational Research, 45:89-125.

OUTCOMES ASSESSMENT AND RE-ACCREDITATION DATA CONCERNS? LOOK TO NATIONAL SURVEYS FOR (SOME) HELP

Marianthi Zikopoulos
Director of Institutional Research
New School University

Introduction

Driven by increased pressure for accountability to government and accrediting agencies, outcomes assessment is currently a major focus in higher education. Our role as institutional researchers in this process is a major one. If not asked to develop the entire plan, with which some of us are faced, we will be, at the very least, asked to provide guidance on the needed data as well as some of the actual data. Institutional researchers with little or no staffing support may want to consider the use of national indicators to fulfill some of the outcomes-assessment-related data needs.

The Outcomes assessment movement

The current emphasis on outcomes assessment is the result of changing views over the past decade and a half about academic quality and effectiveness from an almost exclusive pre-occupation with "**inputs**" and "**processes**" to a more mission-specific focus on outputs and "**outcomes**." In the past, quality was measured by inputs such as the academic preparation of incoming students, quality and reputation of faculty, the number of books and other learning resources available to students, and by processes, such as programs offered, curricular requirements, and availability of student support services. In the 1990's, however, colleges and universities have come under increasing pressure to demonstrate that they provide added "value" to their students, and, in the case of public institutions, contribute to the state economy. The national trend is now toward a results-oriented concern for educational outcomes. Assessment of student learning, in particular, has become the focus.[1]

The emphasis on assessment is driven to a large extent by increased pressure from the federal government, accrediting associations for institutions of higher education, and, more recently, state governments. In 1988, the U.S. Department of Education mandated that accrediting agencies ask institutions to (1) specify their educational goals and (2) conduct student assessment to determine whether they are achieving these goals.2 In response, in the 1990's, the accrediting agencies began requiring that institutions develop a plan "…for assessing their overall effectiveness as well as student learning outcomes."[2] The focus on outcomes assessment has intensified over the decade, with states developing their own mandates for assessment of outcomes.

Currently, our accrediting body, the Middle States Commission on Higher Education (MSCHE), requires that the self-study reports of institutions undergoing re-accreditation include a

comprehensive plan for assessing effectiveness. The Commission specifies two partly overlapping types of assessment:[3]

Assessment of Institutional Effectiveness: evaluation of the institution's a) overall effectiveness in meeting its goals, b) effectiveness in assuring that students achieve the appropriate learning and other outcomes, and c) efficiency in the use of resources.

Assessment of Student Learning: demonstration that students have knowledge, skills and competencies that a) are consistent with the institution's goals and b) meet appropriate higher educational standards.

MSCHE does not prescribe any particular assessment process. It gives institutions much latitude to develop plans that work for them. However, it requires that:

a) assessment address whether the stated goals and objectives of the institution are met

b) that student learning be the primary focus of the assessment process, and

c) institutions demonstrate that they actually use the results of assessment to improve themselves.

Outcomes Assessment for Institutional Effectiveness

Assessment of outcomes, while driven by the need to meet accreditation requirements, can be a powerful tool in enhancing the effectiveness of colleges and universities. If it is built into a strategic planning process, assessment can provide information on how effectively the institution is meeting its goals and point to areas in need of improvement. Even in the absence of strategic planning, if done thoughtfully, assessment can lead to institutional improvement by informing policy decisions on issues of importance.

Put simply, student assessment that is integrated into planning consists of asking:[4]

- *What do we (faculty, deans, administrators) want our students to learn and in how do we want them to grow?*

  Develop clear statements in the institution's **mission, goals, and objectives** regarding desired student outcomes.

- *How will we get there?*

  Plan--what processes are in place for achieving our goals? (resources, curriculum, instruction, student support services, co-curricular activities, etc.)

- *How do we know that we have accomplished what we set out to accomplish?*

Assessment--establish indicators of learning and growing, set benchmarks, establish measures for the indicators; decide on methodology; collect and analyze data.

- *What changes do we need to make, based on the findings of the evaluation?*

  Adjustments to the plan (if necessary)--make appropriate changes in plan that will lead to improvements.

Steps in the Development of Assessment Plans

1. Development of clear statements regarding what we try to accomplish *(mission)*.
2. *Definition of major goals and objectives* regarding student learning.

- What do we want our graduates to know as a result of their collegiate experiences? (in-depth knowledge in discipline; general knowledge in core disciplines)
- What skills do we want them to have? (General education skills, such as critical analysis and reasoning, oral and written communication, scientific and quantitative reasoning, technological competency, and information literacy)
- How do we want them to grow socially and psychologically? (Attitudes and values, such as tolerance for diversity, working effectively as a member of a team, making value-based judgements).

3. Identification of *indicators* (and measures) to assess progress toward achieving the goals.
4. Identification of *methods* for assessing student achievement at important stages of the program. Decisions need to be made as to *when* to assess (e.g. upon entrance, during senior year, etc.) and *who* will do the assessing. Questions that must be considered are: what staff, financial, IT support is there for carrying out an effective assessment plan?
5. Determination of process by which the *results* will be *used* for institutional improvement.
6. Development of *timetable* for accomplishing the previous steps.
7. *Implementation* of assessment plan and *revising* as needed.

Indicators of Student Outcomes

Student learning and growth can be measured indirectly or indirectly. Many indicators are developed by the institution, either centrally or by individual departments. In addition, a number of national indicators exist, which can be used for measuring student outcomes. Advantages of the national measures include time-savings, since they are administered by the developers, and comparative data for a peer or national group. Disadvantages include their generic nature, which may not address the needs of a particular institution and the possibility of low response rates.

Direct indicators include:

- Tests and examinations (local/faculty designed, and commercially produced standardized tests)

- Pre-test/post-test evaluation
- Capstone course evaluation
- Course-embedded assessment
- Portfolio evaluation
- Thesis evaluation
- Videotape and Audiotape evaluation of art exhibit/performance.

Indirect indicators include:

- Curriculum and syllabus analysis
- Student surveys (local and national)
- Exit interviewing
- Alumni/ae surveys
- Employer surveys
- External reviewers

Timing of Assessment

Assessment at important points in time during and after college can provide useful information for assessing program effects on student outcomes or making program improvements. These stages include[5]:

College Entry (Beginning of First Year)--Data collected when students enter college provide a baseline (pre-test) for making meaningful comparisons with data collected at later stages in order to determine whether change has occurred.

Data collection point #2: End of First Year--Assessment at this point may provide feedback on the effectiveness of specific programs designed to enhance learning during the first year of college. It may also provide information on perceptions of students who decide to return and those who do not at this critical point in student retention.

Data collection point #3: End of Sophomore Year/Beginning of Junior Year--To assess progress in student learning and growing.

Data collection point #4: End of Senior Year--To evaluate the extent to which our goals regarding student growth have been achieved.

Data collection point #5: After Graduation--To evaluate the extent to which our goals regarding student growth have been achieved; to assess alumni achievements.

National Instruments that Provide Outcomes Data[6]

Surveys of Entering Students:

CIRP (Cooperative Institutional Research Program) *Freshman Survey*--Developed by UCLA's Higher Education Research Institute (HERI)--administered for more than 30 years, it is

considered the primary indicator of incoming college student attitudes, expectations, and pre-college experiences.

College Student Expectations Questionnaire (CSXQ)--administered by the Indiana University Center for Postsecondary Research and Planning (CPRP), this instrument collects information on student's expectations for their future educational experiences. Can be used with the CSEQ (see below) to track changes.

Surveys of Enrolled Undergraduates

Your First College Year (YFCY)--developed by HERI is a follow-up to the CIRP freshman survey. It also assesses students' experiences with first-year programs such as learning communities, and introductory courses. Can be used as baseline for longitudinal follow-up with CSS survey (see below) to assess students' self reported skill development and changes in attitudes and behaviors.

College Student Survey (CSS)-- developed by HERI--assesses incoming students' expectations--Measures students' experiences and satisfaction to assess changes since entry to college (used as a follow-up to CIRP)

College Student Experiences Questionnaire (CSEQ)--Administered by the University of Indiana--Evaluates quality of students' in-class and out-of class experiences, progress toward learning and personal development goals, satisfaction, and perceptions of the environment.

National Survey of Student Engagement (NSSE)--Developed by a panel of leading assessment scholars as a model for quality of undergraduate education. Administered by the University of Indiana, it collects data on students' engagement in effective educational practices (level of challenge, active learning, student-faculty interaction, supportive environment, etc.)

Student Satisfaction Inventory (SSI)--Developed by Noel-Levitz--measures student satisfaction with compared to perceived importance of various aspects of their college experience.

Student Proficiencies and Learning Outcomes

Collegiate Assessment of Academic Proficiency (CAAP)--Developed by ACT, it assesses students' achievement in general education skills.

Academic Profile--Developed by ETS and The College Board--measures general education skills.

Tasks in Critical Thinking--Assesses proficiency in college-level higher order thinking skills.

Major Field Tests--Created by ETS, they measure students' academic achievement in major fields of study.

Alumni/ae Surveys

Comprehensive Alumni Assessment Survey (CAAS)--Administered by the National Center for Higher Education Management Systems (NCHEMS). It assesses institutional effectiveness and provides data on alumni/ae personal development and career preparation.

College Results Survey (CRS)--Administered by Petersons. Identifies values, abilities, work skills, and participation in life-long learning of college graduates.

POSSIBLE INDICATORS FOR MAJOR STUDENT LEARNING OUTCOMES[7]

GOAL	INDIRECT	DIRECT
General Education	Local: Surveys of entering and enrolled seniors and alumni/ae National: CIRP--Freshman Survey CSXQ Survey HERI--CSS CSEQ NSSE	Local: Student grades in General Education courses Course-embedded assessments National: CAAP Academic Profile Tasks in Critical Thinking
Knowledge in the area of concentration	Senior surveys assessing education in the major Alumni/ae surveys assessing education in the major	Local: Comprehensive exam Course-embedded assessment Capstone course Thesis/Research project Performance/Exhibit Internship/Field Work Portfolio National: Major Field Tests
Personal/Social Growth Maturity & Satisfaction	Local: Senior Surveys Student Satisfaction Surveys National: CIRP and CSS CSXQ and CSEQ NSSE SSI	Local: Interviews Observations Focus Groups

Outcomes assessment is here to stay (for a long while it seems). Our role as institutional researchers is a major one. If not asked to develop the entire plan, with which some of us are faced, we will be asked to provide guidance on the needed data as well as some of the actual data. Institutional researchers with little or no staffing support may want to consider the use of national indicators to fulfill their data needs.

References

[1] Peterson, M. L., Augustine, C. H., & Vaughan, D.V. (1999). *Designing Student Assessment to Strengthen Institutional Performance in Baccalaureate Institutions*. Stanford, CA: Stanford University, National Center for Post-Secondary Improvement.

[2] Middle States Commission on Higher Education (1994). *Characteristics of Excellence in Higher Education.* (p. 16) Philadelphia, PA: Middle States Commission on Higher Education.

[3] Middle States Commission on Higher Education (1994). *Characteristics of Excellence in Higher Education.* (p. 16) and Middle States Association Commission on Higher Education (2001) *Characteristics of Excellence in Higher Education: Standards for Excellence.* (Draft for Discussion), (pp. 18, 42), Philadelphia, PA: Middle States Commission on Higher Education.

[4] Adapted from Alan Sturz: "A User's Guide to Institutional Effectiveness." Paper presented at the North East Association for Institutional Research conference, Philadelphia, PA, November 1998.

[5] Cuseo, Joseph B. (2000). "Assessment of the First Year Experience: Six Significant Questions". (pp.4-6) ~http://brevard.edu/fyc/CuseoRemarks.htm.

[6] The information presented in this section was obtained from Borden, V. H. and Owens, J. L. (2001). *Measuring Quality: Choosing Among Surveys and Other Assessments of College Quality.* Washington, D.C.: American Council on Education.

[7] Adapted from Volkwein, J. Fredericks. (2001) *Responding to Accreditation and Assessment on Your Campus: Why, What, Who, How?* Workshop presented at NEAIR conference, Boston MA, Nov. 17-20.