USE OF COURSE EVALUATIONS TO ASSESS THE CONTRIBUTIONS OF CURRICULAR AND PEDAGOGICAL INITIATIVES TO UNDERGRADUATE GENERAL EDUCATION LEARNING OBJECTIVES

ROBERT J. THOMPSON JR. AND MATT SERRA
Duke University

This report describes measures and methods of data analysis and representation of a systemic approach to the assessment of curricular and pedagogical contributions to general education learning objectives. An existing system for student evaluation of teaching was transformed to enable faculty to designate the learning objectives of the course and students to evaluate the extent to which the course contributed to their self-appraised growth in intellectual skills. This systemic approach enables analysis at the level of the individual course and across courses with common learning objectives and pedagogical approaches. The utility of this system is demonstrated through assessing the differential contribution to general education learning objectives of curricular components and an innovative pedagogical approach that blends field-based research and service learning.

Motivated by recognition of the inadequate state of undergraduate education, research universities are engaged in efforts to renew undergraduate education through curricular and pedagogical initiatives. There is an emphasis on general education, curricular coherency, and pedagogies of engagement. Correspondingly, the need for assessment of the contributions of curricular and pedagogical initiatives to student learning outcomes is well recognized (Chun, 2002).

Developing methods and procedures to assess student learning outcomes presents a set of challenges. A curriculum has multiple objectives and is made up of a myriad of individual courses; multiple pedagogies are employed; and students do not take the same courses in the same sequence. Given this scope and variability, systemic approaches that are comprehensive but yet cost-effective are needed.

Most institutions of higher education have developed at least one systemic measure: Student Evaluations of Teaching (SET). An extensive research literature has been generated regarding the psychometric properties, factor structure, and validity of measures of instructional effectiveness (díApollonia & Abrami, 1997). The purposes of student ratings include providing feedback to instructors to enhance their effectiveness, informing student selection of courses, providing information for faculty promotion review (Feldman, 1997) and providing process and outcome measures that can be used for research (Marsh & Dunkin, 1997).

It was recognized that the traditional SET measure could be modified to go

beyond the focus on satisfaction with teaching and description of instructional dynamics to also address the student learning objectives of the institution. That is, the traditional SET can be transformed to enable faculty to designate the learning objectives of the course and students to evaluate the extent to which the course contributed to their attainment of specific learning objectives. Such a transformation would enable assessment of the curriculum and pedagogical initiatives through linkage to an existing system and infrastructure and add another purpose to the traditional SET: To repeatedly inform the students about the learning objectives of the curriculum (Ratcliff, 2003).

This systemic approach relies on faculty and student self-report data and would need to be just one component of an overall assessment plan that also included direct measures of student learning. However, this system is in accord with the literature on conditions under which the validity of self-reports is enhanced (Kuh, 2001). Furthermore, faculty intentions and student self-appraisals are important to assess in their own right and the system affords direct measures of these phenomena.

Duke University has been engaged in the process of enhancing undergraduate education within Trinity College of Arts and Sciences. The formulation of the new general education curriculum was guided by the articulation of an intellectual agenda for undergraduate education that delineated learning objectives in terms of intellectual skills and broad dispositions viewed as necessary for functioning in leadership roles in the twenty-first century (Table 1). In addition, we have encouraged pedagogies that incorporate experiential learning, such as service learning. Concurrently, we revised the student course evaluation process and aligned it with the learning objectives to provide a systemic measure of learning outcome: Student self-appraised growth in general intellectual skills. The purpose of this article is to demonstrate how course evaluations can be utilized to assess the differential contribution of curricular and pedagogical initiatives to self-appraised student learning outcomes. First, we describe the intellectual agenda, new curriculum, and a specific pedagogy of engagement that provide the context for the assessment of learning outcomes.

Intellectual Agenda

A review of the undergraduate curriculum in Trinity College of Arts and Sciences resulted in a reaffirmation of Duke's long standing tradition of providing a liberal education that fosters development of intellectual skills, in particular, the ability to think critically, rationally, and in historically and ethically informed ways; to analyze, evaluate, and synthesize information; to problem solve; and to communicate effectively. As a research university, the common bond of all members of the Duke academic community is the pursuit of increased knowledge and understanding through the processes of inquiry, discovery, and synthesis. Undergraduates, as members of the academic community, are expected to participate in the discovery and learning processes. Thus, we intend for the undergraduate experience to be inquiry-based and for our students to be epistemologically sophisti-

cated, life-long learners who have an understanding of how knowledge is generated, organized, accessed, and applied to pressing social needs. We seek to foster fluency across the boundaries of knowledge and across cultural boundaries. Overall, we view a quality education as enabling students to bring meaning to information and to discern among competing claims.

Architecture of The Curriculum

A matrix approach (Figure 1) was adopted as the organizational structure to reflect learning objectives of the new curriculum in terms of two dimensions: Areas of Knowledge and Modes of Inquiry. Undergraduate Courses at Duke have historically been categorized in terms of Areas of Knowledge reflecting both differences in subject matter and methods of discovery. Five areas of knowledge serve as the vertical axis of the matrix. Six Modes of Inquiry consistent with the learning objectives of our new curriculum constitute the horizontal axis of the matrix: Cross Cultural Inquiry; Science Technology and Society; Ethical Inquiry; Foreign Language; Research; and Writing, compromised of a first year academic writing course (W20) and two subsequent writing in the discipline courses (WID). In addition to addressing learning objectives in terms of general intellectual skills and broad dispositions, more fine-grained learning objectives were delineated for each mode of inquiry. A single course can simultaneously address a substantive topic or area of knowledge and also teach systems of reasoning or modes of inquiry. Furthermore, upper level courses in the major can also simultaneously contribute to general education learning objectives.

Pedagogy of Engagement

Research service learning is a blend of two pedagogies of engagement: Field-based research and service learning. The service that students provide is field-based research on an issue or problem that is of importance to the community partner. Through reflection and analyses, students integrate their research project with the objectives and content of an academic course.

Measures

Concurrent with the advent of our new curriculum, we developed a new Student Course Evaluation (SCE) instrument and a Faculty Course Description form (FCD). With the SCE, students rate on a scale of 1 (low) to 5 (high) the quality of the course and the quality of instruction; characteristics of the course in terms of amount of work, difficulty, and intellectual stimulation; course dynamics; and the extent to which the course contributed to their growth along several dimensions of intellectual skills, broadly based on Bloom's (1984) taxonomy. Within each section students are provided with ample room for written comments. In particular, the course characteristics and intellectual growth dimensions provide a rich source of data for evaluating individual courses but also curricular components, pedagogical approaches, and programmatic initiatives. With the FCD, faculty characterize the learning objectives of the course, along the same dimensions of intellectual skills as the SCE, pedagogical approach(es), and use

of instructional technology.

Procedures

The SCE and FCD forms are distributed each semester through the Trinity College Office of Assessment (TCOA) to the Director of Undergraduate Studies (DUS) in each of the 35 departments and programs offering undergraduate courses in Trinity College. SCE forms are completed anonymously by students at the end of the semester and both the SCE and FCD forms are returned to the TCOA and scanned into a relational database. Each Fall and Spring semesters, the TCOA receives approximately 18,000 SCE and 900 FCD forms across approximately 1,100 courses. A report is prepared for each course that provides the means and distributions of ratings for each item and the corresponding means across all the courses offered that term in the College. In addition, each department is provided a composite report for all courses offered in the department that term. Faculty are provided a report of each of their courses and department chairs use the data in evaluating faculty teaching in conjunction with promotion reviews.

Analysis of Learning Outcomes:

With the SCE in a relational database, comparisons can be accomplished of individual courses, clusters of courses, or specific programs by transforming raw scores into "Z"scores with a mean of zero and a standard deviation of one based on the data for all the courses offered in the College that term. Comparisons between a course or types of courses can be visually represented in terms of Z score deviations above and below the mean for the College.

To demonstrate the use of this method to evaluate the "value added" of components of our new curriculum, we present the SCE data for courses with Ethical Inquiry (EI), Science Technology, and Society (STS), Academic Writing (W20), and Research (R) curricular designations offered in Fall 2001 and Spring 2002. Tables 2-5 present the specific learning objectives for these courses. In addition, we present findings for courses that incorporated a research service learning (RSL) pedagogy.

Findings

Figure 2 presents the SCE data of courses with Modes of Inquiry designations as Science, Technology, and Society (STS) or Ethical Inquiry (EI) compared with the average for all courses in Trinity College as the zero point. In terms of course characteristics, students rated the quality and intellectual stimulation of the EI courses higher than STS courses and more than a quarter of a standard deviation above the mean for all courses in the College. In terms of contribution to intellectual skill development, the relative higher "value added" of EI courses compared with the average for all courses in the College courses and contrasted with STS courses is clear.

Figure 3 presents the SCE data for the first year writing course (W20), and Research Courses (R). Since the W20 courses are small seminars, this contrast was controlled for class size of < 19. That is, only seminars were used in the com-

putation of the average for the College. The first year writing course is rated comparably to other seminars in terms of quality and difficulty but higher in amount of work and intellectual stimulation. Research courses are rated higher in quality and stimulation but comparably in difficulty to the average for the courses in the College. The differential "value added" in terms of self-appraised development of intellectual skills by these courses is apparent. W20 courses are not rated notably higher in contributing to gaining factual knowledge or understanding fundamental principles but are rated high in contribution to development of analysis, evaluation, and writing skills. The Research courses are rated higher in contribution to development of intellectual skills than the average for all the courses in the College with the exception of evaluating the merits of ideas and oral impressions.

Figure 4 presents the SCE data for research service learning (RSL) courses. These courses are rated a quarter of a standard deviation higher in quality and intellectual stimulation than the average for the College and lower in amount of effort and difficulty. In particular, the courses are rated higher in the contribution to intellectual skills of applying concepts and synthesizing knowledge that are the specific objectives of RSL courses.

The FCD forms enable the faculty to characterize how essential the development of specific intellectual skills are to the objectives of the course. Figure 5 presents the comparisons of the percentage of faculty rating development of specific intellectual skills as essential across courses with R, EI, RSL, and W20 designations. It can be seen that faculty characterize the learning objectives differently across these curricular and pedagogical components of undergraduate education. For example, faculty generally do not consider "gaining factual knowledge" as an essential learning objective for the first year writing course (W20) but more than 90% see "learning to analyze" as an essential learning objective for these courses. Similarly, Ethical Inquiry courses are noteworthy because 86% of the respondents characterized "understanding fundamental concepts and principles" as an essential learning objective of these types of courses.

Discussion

Using course evaluations as a method of assessing student learning has several strengths. First, the approach is systemic. It builds on a well established tradition and system at Duke for students to rate each of their courses each term. The content of the process was modified to align it with the intellectual agenda for undergraduate education. Second, centralizing the scoring and reporting functions through the Trinity College Office of Assessment relieved individual departments of these tasks and the information provided to individual faculty and departments is more extensive and enables comparison with salient referent groups. For example, the ratings for an individual course are compared to the composite ratings for all the courses in the department and the composite for the department can be compared with the composite for all courses within a division, such as the social sciences or humanities. Third, having the data in a central rela-

tional database enables appropriately controlled analyses that can guide refinements of curricular and pedagogical initiatives. For example, the findings reported above supported the "value added" of R, EI, and W20 components of our new curriculum but the STS courses were not rated highly in terms of intellectual stimulation or contribution to intellectual growth. We subsequently have focused on strengthening the STS courses.

These analyses also contribute to the question of the validity of student appraisals of teaching and learning (Boud, 1995). First, the findings across the array of analyses indicate that students make meaningful differentiations. For example, individual courses are not rated uniformly across all dimensions and courses with different objectives are rated differently on specific dimensions. Furthermore, the differentiations that students make have face validity. For example, courses that faculty rate low in terms of gaining factual knowledge as an essential learning objective are generally not rated high by students in contribution to this dimension of intellectual growth. Second, the findings do not support the frequent contention that student appraisal are primarily a function of the ease of the course, that is, that appraisals are most strongly correlated, negatively, with perceived degree of difficulty of the course or amount of work required. Our data indicate that ratings of the quality of the course are most strongly correlated, positively, with appraisals of the intellectual stimulation of the course ($r = .57$) and very weakly with difficulty ($r = .07$) and amount of work ($r = .13$). Thus, the findings provide support for the utility of student course evaluations for appraisal of teaching and learning. However, we also consistently make the point that student appraisal are only one measure of teaching and learning and that the full assessment portfolio needs to include direct measures of teaching and learning outcomes.

References

Bloom, B.S. (Ed.). (1984). Taxonomy of educational objectives Book 1: Cognitive Domain. White Plains, NY: Longman.

Boud, D. (1995). Enhancing learning through self-assessment. London: Kogan Page.

Chun, M (200). Looking where the light is better: A review of the literature on assessing higher education quality. Peer Review, Winter/Spring pp. 16-25.

d'Apollonia, S. & Abrami, P.C. (1997). Navigating student ratings of instruction. American Psychologist, 52, 1198-1208.

Feldman, K.A. (1997). Identifying exemplary teachers and teaching: Evidence from student ratings. In R.P. Perry & J.C. Smart (Eds.) Effective teaching in higher education: Research and practice (pp.368-395). New York: Agathon Press.

Kuh, G.D. (2001). Assessing what really matters to students. Change, 33, pp. 10-17.

Marsh, H.W. & Dunkin, M.J. (1997). Student's evaluations of university teaching: A multidimensional perspective. In R.P. Perry and J.C. Smart (Eds.). Effective teaching in higher education: Research and practice (pp. 241-320). New York: Agathon Press.

Ratcliff, J. L. (2003). Dynamic and Communicative Aspects of Quality Assurance. Quality in Higher Education, 9, 117-131.

Table 1
Student Learning Objective

- Intellectual Skills
  - Critical thinking and reasoning
  - Critical reading
  - Formulating, supporting, and evaluating an argument
  - Problem solving
  - Analyzing, integrating, and synthesizing information and ideas
  - Writing effectively
- Broad Disposition/Understandings
  - Epistemologically sophisticated
    - o Knowledge about knowledge
    - o Fluency across the boundaries of knowledge
    - o Integration of knowledge across disciplines
  - Cross cultural fluency
  - Scientific and quantitative literacy
  - Civic and social responsibility (active agency for community change)
  - Collaboration
  - Life-long learning

Table 2
Learning Objectives for Science, Technology, and Society (STS) Courses

- Understanding of the historical, social, political, and/or economic roots of scientific or technological fields or phenomena

- Understanding of contemporary issues relating to the development and application of a particular area of science and technology

Table 3
Learning Objectives for Ethical inquiry (EI) Courses

- Articulate ethical questions regarding individual and social behavior, institutions, and ways of life
- Discern and choose among competing claims, distinct systems of values, and courses of action

Table 4
Learning Objectives for Research Intensive Courses

- Formulate a question, analyze material, and integrate findings
- Participate in a mentoring relationship with faculty
- Develop a research paper, poster session, performance, or product that describes or exemplifies an understanding of how knowledge in the discipline is generated, organized, and presented

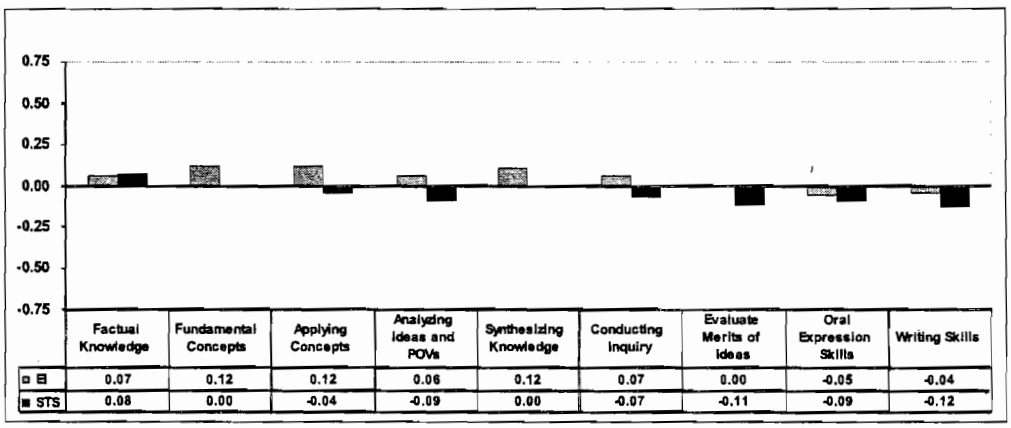
Table 5
Learning Objectives for Writing Courses

- Ability to read in a scholarly and critical fashion and to distinguish between expressive or poetic and analytic argumentative forms of writing
- Ability to analyze, integrate, and synthesize information and ideas
- Ability to develop, support, critique, and refine arguments
- Ability to write clear and engaging text, attending to conventions and style appropriate to audience and purpose

Figure 1. Trinity College Curriculum

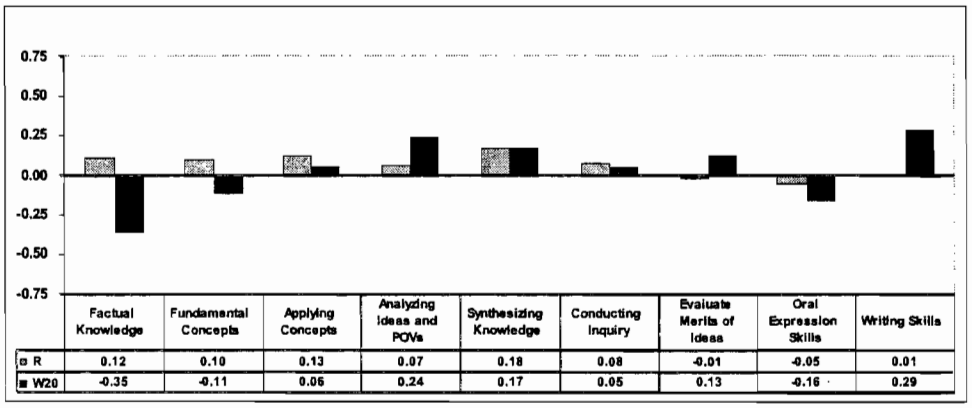
	Modes Of nquiry					
Areas of Know edge	Cross Cu tura nquiry	Science, Techno ogy & Society	Ethica nquiry	Foreign Language	Writing	Research
Arts, Literatures, and Performance (2)						
Civi izations (2)						
Socia Sciences (2)						
Natura Sciences (2)						
Quantitative Studies (2)						
(Minimum Required)	(2)	(2)	(2)	1 (up to 2 more)	(3)	(2)

Figure 2.
Mean Z-Scores for Ethical Inquiry (EI) and Science, Technology, and Society (STS) Courses vs. Trinity Overall



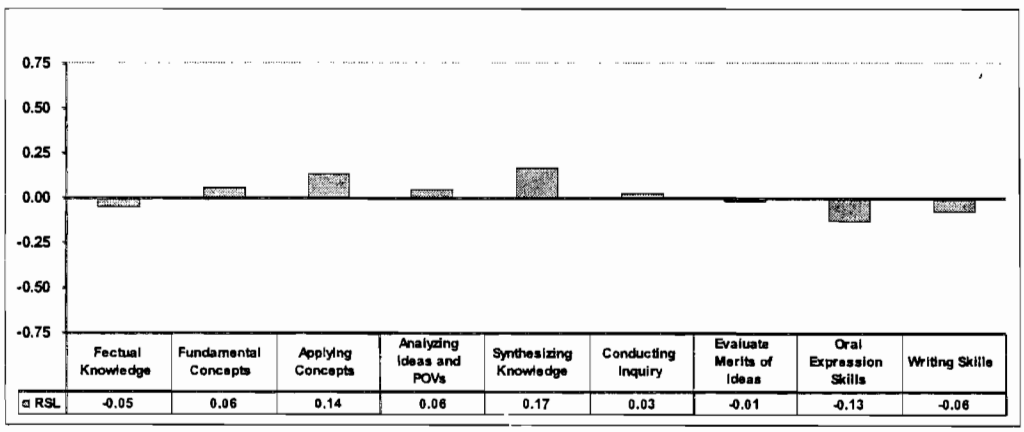
	Factual Knowledge	Fundamental Concepts	Applying Concepts	Analyzing Ideas and POVs	Synthesizing Knowledge	Conducting Inquiry	Evaluate Merits of Ideas	Oral Expression Skills	Writing Skills
EI	0.07	0.12	0.12	0.06	0.12	0.07	0.00	-0.05	-0.04
STS	0.08	0.00	-0.04	-0.09	0.00	-0.07	-0.11	-0.09	-0.12

Figure 3.
Mean Z-Scores for Research (R) and Academic Writing (W20) Courses vs. Trinity Overall



	Factual Knowledge	Fundamental Concepts	Applying Concepts	Analyzing Ideas and POVs	Synthesizing Knowledge	Conducting Inquiry	Evaluate Merits of Ideas	Oral Expression Skills	Writing Skills
R	0.12	0.10	0.13	0.07	0.18	0.08	-0.01	-0.05	0.01
W20	-0.35	-0.11	0.06	0.24	0.17	0.05	0.13	-0.16	0.29

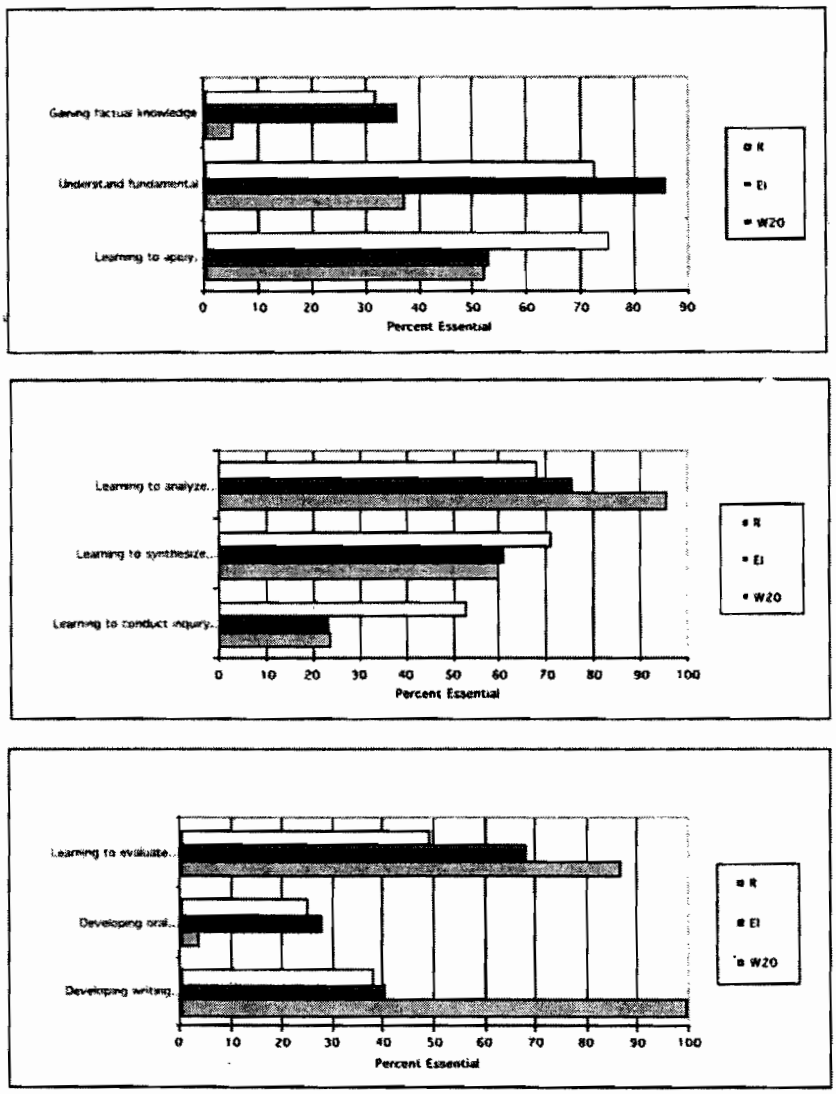
Figure 4.
Service Learning (RSL) Courses vs. Trinity Overall



	Fectual Knowledge	Fundamental Concepts	Applying Concepts	Analyzing Ideas and POVs	Synthesizing Knowledge	Conducting Inquiry	Evaluate Merits of Ideas	Oral Expression Skills	Writing Skills
RSL	-0.05	0.06	0.14	0.06	0.17	0.03	-0.01	-0.13	-0.06

Figure 5.

Percentage of Faculty Rating Learning Objectives as Essential for Research (R), Ethical Inquiry (EI), Research Service Learning (RSL), and Academic Writing (W20) Courses



This article was reprinted from Education
Volume 125 Summer/2005 No. 4

The figures (following the article) were omitted in error in the original publication.