IDEA Technical Report No. 18

Revising the IDEA Student Ratings of Instruction System 2002-2011 Data

Stephen L. Benton
Dan Li
Ron Brown
Meixi Guo
Patricia Sullivan

February, 2015

Table of Contents

List of Tables and Figures

Introduction

This report describes the processes undertaken to revise the IDEA Student Ratings of Instruction (SRI) system. The previous revision occurred in 1999, as described in IDEA Technical Report No. 11, *Revising the IDEA System for Obtaining Student Ratings of Instructors and Courses* (Hoyt, Chen, Pallett, & Gross, 1999). The procedures explained herein occurred across a three-year period, beginning in the Spring of 2011 and ending with the development of pilot instruments in the Fall of 2014.

Since its inception, the purpose of IDEA SRI has been to improve instruction. Donald P. Hoyt (1973a) developed the original system in 1969 with the help of faculty and students. Over the next few years, revisions were made based on input from users, colleagues, and experts in the fields of educational measurement and teaching. In 1975, with support from the Kellogg Foundation, a version of IDEA SRI was developed that remained largely unchanged for the next 24 years (see Hoyt & Cashin, 1977, for a description of the process). Since then, three core beliefs have distinguished IDEA SRI from other student ratings systems:

- The chief measure of teaching effectiveness is the *amount of progress students make on learning objectives stressed by the instructor*.
- Student-learning outcomes should reflect the purpose of instruction.
- A given teaching method might be more effective with certain learning objectives than with others.

The need for the current revision seemed obvious considering the many changes that have occurred in higher education since 1999. The increase in the number of courses offered online is but one example. Another is the increasing ubiquity of mobile devices, which enable quick assessment and feedback. In addition, accreditation standards have elevated student-learning outcomes not emphasized in 1999, notably civic engagement, ethical reasoning, diversity, global awareness, and quantitative literacy. Moreover, the nature of faculty appointments has changed. Part-time and adjunct faculty account for nearly half of all faculty (excluding graduate employees) and for almost 70% of instructors in community colleges (American Federation of Teachers, 2010).

In light of these changes, we embarked on an update of IDEA SRI that would take advantage of mobile technology, incorporate contemporary learning outcomes, and provide helpful feedback to both full- and part-time faculty. Throughout, we sought to retain certain key features:

- A focus on relationships between teaching methods and measures of teaching effectiveness
- Statistical control of extraneous factors that influence student ratings but are not under the instructor's control
- A focus on instructional improvement

We turned to multiple information sources to guide the revision: extensive statistical analyses of the IDEA SRI research database; the professional literature on teaching and learning;

accreditation criteria from national and regional institutional accrediting agencies; the Association of American Colleges and Universities (AAC&U) VALUE rubrics; IDEA staff; focus groups of IDEA users; expert panels comprised of psychometricians, statisticians, higher-education administrators, and experts in faculty development and evaluation; university faculty; and college students. Every effort was made to ensure an inclusive process.

The procedures described in this report ultimately led to the modification of the existing IDEA student ratings instruments as well as the creation of a new one. The proposed instruments and their purposes are described in the final section of the report.

The Revision Process

Our revision effort started with focus groups held at two professional conferences to collect feedback from veteran IDEA users. We then surveyed IDEA staff to obtain their suggestions. An updating team including staff members from relevant units at IDEA was formed to guide the revision process. Two expert panels were recruited to review drafts of revisions multiple times and provide valuable feedback. The updating team made further revisions to incorporate experts' suggestions. We then contacted IDEA users and non-users for additional comments on the updated revisions. Finally, we conducted cognitive interviews with more than 20 college students to test our proposed revisions.

Focus Groups

To collect feedback from users' perspectives, we conducted two focus groups early in the updating process. The first occurred at IDEA's Users Group Meeting held in Nashville, TN, in October 2012. Twenty current users of IDEA participated. The second was held at the Train the Trainer Workshop in San Antonio, TX, in February 2013. It was comprised of 12 individuals who were veteran users of IDEA SRI. Both groups were of mixed gender and ethnic diversity, and included representatives from public and private institutions of various sizes.

To obtain feedback on IDEA SRI learning objectives, teaching methods, as well as course and student characteristics, we used three questions to guide the discussion: (a) Which items are no longer relevant? (b) What items are missing? and (c)What are other major issues and questions should be considered? After analyzing notes and transcripts, the following salient themes emerged from the focus groups:

Learning Objectives

Although focus group participants agreed that all existing learning objectives were appropriate, they suggested two pairs of learning objectives could be combined due to their conceptual similarity. Since Objectives 1 (Gaining factual knowledge [terminology, classifications, methods, trends]) and 2 (Learning fundamental principles, generalizations, or theories) both pertain to cognitive learning outcomes (Hoyt & Cashin, 1977), instructors tend to emphasize them simultaneously in a course. The other pair of objectives that potentially could be combined are 9 (Learning how to find and use resources for answering questions or solving problems) and 12 (Acquiring an interest in learning more by asking my own questions and

seeking answers). Both items tap into lifelong learning (Hoyt et al., 1999), with one focusing on developing practical skills and the other emphasizing cultivating the disposition. Since both elements are central to encouraging lifelong learning, they might be combined into a single learning objective.

Other issues raised pertained to wording of items. Participants pointed out that due to the order of examples in Objective 6 (Developing creative capacities [writing, inventing, designing, performing in art, music, drama, etc.]), instructors in some disciplines (e.g., engineering, science) tend to consider the objective being primarily about creativity in writing and arts and inapplicable for their fields (e.g., inventing, designing). In addition, although Objective 9 captures the two important steps in information use, it should include "evaluate" to address an essential aspect of information literacy.

In terms of potential additional objectives, participants suggested the following topics:

- Civic engagement
- Ethical decision making
- Global awareness and diversity
- Quantitative literacy
- Research

Teaching Methods

Comments from participants focused on four items: Methods 3, 7, 9, and 20. Some suggested that Method 3, *Scheduled course work (class activities, tests, projects) in ways which encouraged students to stay up-to-date in their work*, lacks relevance and could be eliminated. Others believed the word "criticisms" in Method 7 (*Explained the reasons for criticisms of students' academic performance*) should be replaced with a neutral word. Also, Methods 7 (above) and 17 (*Provided timely and frequent feedback on test, reports, projects, etc. to help students improve*) both address feedback given to students and could be candidates for combination. Method 9, *Encouraged students to use multiple resources (e.g., data banks, library holdings, outside experts) to improve understanding*, contains obsolete terms and needs updating. Participants also pointed out that Method 20, *Encouraged student-faculty interaction outside of class (office visits, phone calls, e-mail, etc.)*, is problematic due to its limited applicability in certain course circumstances (i.e., online and hybrid courses).

When asked to consider teaching methods not included in the current instrument, participants suggested the addition of a method for "encouraging student self-reflection."

Course and Student Characteristics

Participants suggested the addition of a series of questions to address coursework: amount of reading assigned, amount of reading I did, and amount of writing.

Interviews with IDEA Staff

In December 2012, individual interviews were conducted with IDEA staff members to solicit ideas about changes they would like to see made to the student ratings system. We considered staff input of great importance because they all worked in the domain of student ratings. Some suggestions aligned with what we learned from focus groups, further emphasizing the needs for those revisions. Unique contributions from those interviews can be summarized as follows:

Learning Objectives

Although Objective 10, *developing a clearer understanding of, and commitment to, personal values*, is conceptually related to ethical decision-making processes, it is not explicitly so. Moreover, personal values are open to interpretation and do not necessarily denote positive views. This item should be replaced with one addressing ethical reasoning.

Teaching Methods

IDEA staff suggested changes to the following items:

- Method 5: Formed "teams" or "discussion groups" to facilitate learning. Staff suggested the removal of the quotation marks. In addition, because the item intends to measure group work, discussion groups may be too limiting and should be changed to *groups*.
- Method 14: Involved students in "hands on" projects such as research, case studies, or "real life" activities. Similar to Method 5, the quotation marks are unnecessary and should be removed.

Experimental Items

The following four items were created for experimental purposes. Research staff have collected adequate data from those items, and they can be removed to shorten the instrument.

44. The instructor used a variety of methods—not only tests—to evaluate student progress on course objectives.

45. The instructor expected students to take their share of responsibility for learning.

46. The instructor had high achievement standards in this class.

47. The instructor used educational technology (e.g., Internet, e-mail, computer exercises, multi-media presentations, etc.) to promote learning.

IDEA SRI Updating Team

In January of 2013, the IDEA SRI Updating Team was formed, comprised of key individuals representing all aspects of the IDEA student ratings system.[1] The purpose of the team was to guide revision planning and to act as a sounding board for proposed changes. The team met regularly, provided substantial input into the revision process, and made invaluable contributions to the revised and new IDEA SRI instruments.

Review of the Literature

In preparing for this revision, IDEA staff reviewed literature pertaining to teaching, learning, and faculty development to identify potential topics for new items. They kept abreast of current issues in the field of higher education through such resources as *The Chronicle of Higher Education, Inside Higher Education, Academic Impressions,* the Professional and Organizational Development (POD) Network's annual *To Improve the Academy*, and the American Association of Colleges and Universities' (AAC&U) *Diversity & Democracy* and *Peer Review*. Staff members also reviewed the AAC&U VALUE rubrics and multiple accreditation standards. Textbooks by the following authors were also helpful: Finkel (2000), Hativa (2013a, 2013b), Svinicki and McKeachie (2011), and Schwartz and Gurung (2012). In addition, members of the IDEA research unit published comprehensive reviews of the literature on student ratings (i.e., Benton & Cashin, 2011; Benton & Cashin, 2014). Based on these sources, the following topics were identified as potential additions:

Learning Objectives

- Developing an understanding of diverse perspectives
- Developing ethical decision-making skills
- Developing global awareness
- Fostering civic engagement
- Developing quantitative literacy/reasoning
- Appreciating cultural diversity

Teaching Methods

- Encouraging student self-reflection
- Facilitating diverse perspectives

Student Characteristics

- Self-efficacy

Expert Panels

In Spring 2013, the first of two expert panels was formed for the purpose of obtaining feedback about proposed changes to the system. The 15 panel members included experts in teaching and learning, technology, measurement, faculty development, faculty evaluation, higher education administration, and institutional assessment (see Appendix G for the identities of the

[1] We are grateful to Ron Brown, Shelley Chapman, Christopher Conner, Sally Garvin, Jake Glover, Tanner Ratzlaff, Ken Ryalls, Angela Simons, Pat Sullivan, and Todd Wallentine who served on the Updating Team.

experts). Based on external and internal feedback as well as literature, the updating team created the first draft of proposed revisions and solicited feedback from the expert panel members. The panel provided feedback to the updating team, which then met to review the input and to modify proposed changes. The panel then reviewed the second draft of proposed changes. After two rounds of review and modification, followed by deliberations with the updating team, we created a draft of proposed revisions for further review.

In Summer 2013, we recruited the second expert panel to gain fresh and diverse perspectives on the proposed revisions, which incorporated feedback from the first expert panel. Similarly, the panel was composed of experts from various aspects of higher education (see Appendix G for the identifies of the experts). The second panel reviewed the third draft of revisions and provided feedback, which served as a basis for further revisions. In November 2013, we circulated the fourth draft to both expert panels and finalized the consultation by another iteration of reviewing feedback and making modifications. The proposed revisions at that time, based on feedback from both expert panels, were as follows:

Learning Objectives

Objectives 1 and 2 should be combined into the following single objective: *Gaining a basic understanding of the subject (e.g., factual knowledge, methods, principles, generalizations, theories).*

Minor changes to Objective 6 should be made to improve its applicability for science and engineering disciplines: *Developing creative capacities (inventing; designing; writing; performing in art, music, drama, etc.).*

Objectives 9 and 12 should be combined to represent life-long learning: *Learning how to find, evaluate, and use resources to explore a topic in depth*.

Objective 10 should be eliminated due to its ambiguity and replaced with an objective that directly addresses ethical decision making: *Developing ethical reasoning and/or ethical decision making.*

Objective 12 should be reworded to emphasize life-long learning: *Becoming a self-directed learner—learning how to ask questions and seek answers.*

Moreover, we proposed to add three new objectives to address diverse perspectives, quantitative literacy, and civic engagement:

Developing knowledge and understanding of diverse perspectives, global awareness, or other cultures.

Learning appropriate methods for collecting, analyzing and interpreting numerical information.

Learning to apply knowledge and skills to benefit others or serve the public good.

Teaching Methods

Methods 7 and 17 should be combined into the following single item: *Provided meaningful feedback on students' academic performance.*

Method 9 should be revised to include evaluation, a crucial component of information literacy: *Encouraged students to evaluate and use multiple resources to improve understanding.*

We proposed to add three new teaching methods to address self-reflection, diversity, and service learning:

Encouraged students to reflect on and evaluate what they have learned.

Helped students to interpret subject matter from diverse perspectives (e.g., different cultures, religions, genders, political views, etc.).

Created opportunities where students can apply course content outside the classroom.

Student and Course Characteristics

As the literature suggested, we proposed to add an item for self-efficacy: *When this course began I believed I could master its content.*

Overall Summary Evaluations

We proposed to drop one of the three overall summary evaluation items because, as one astute expert panel member pointed out, it provides ambiguous results: *As a result of taking this course, I have more positive feelings toward this field of study.* A low score on this item could evidence either effective or ineffective teaching, depending on the student's aptitude and interest in the content area.

Other Consultation

We also consulted with several faculty members who taught quantitative methods courses in statistics, engineering, and psychology. They provided invaluable input in crafting the quantitative literacy objective.

In addition, we enlisted the help of a faculty member with expertise in the domains of civic engagement and service learning. She assisted us in adopting the following learning objective and teaching method:

Learning to apply knowledge and skills to benefit others or serve the public good.

Created opportunities for students to apply course content outside the classroom.

Finally, we shared all items with an English professor who provided very helpful editing for readability and clarity.

Feedback on Proposed Revisions from IDEA Users and Non-users

During the revision process, it was critically important to involve users of IDEA SRI. We also intended to have the revised instrument examined by people who had no experience with IDEA. Therefore, we conducted a survey with over 25 faculty members across the country to collect their feedback. The questionnaire was composed of two parts, which focused on the *Diagnostic Form* (DF) and *Faculty Information Form* (FIF), respectively. A mockup instrument including existing items and proposed revisions was presented to the respondents. An open-ended question was presented below each section of the instrument to collect respondents' major concerns and elaborations, if any. To facilitate the identification of revised items, we color-coded revisions based on their item type; new items, significantly modified items, and items with minor revisions were presented in different colors. After the survey ended, we content analyzed responses to the open-ended questions to identify concerns and suggestions for revisions.

Cognitive Interviews with Students

Another strategy employed to evaluate the new and revised items was cognitive interviews conducted with college students. We recruited more than 20 students from a large public university in the Midwest to individually interview. To ensure we obtained diverse views, we stratified by gender, age, academic major, class level, and English proficiency.

During the interviews, which took approximately 30 to 45 minutes each, the researcher showed the participant a handout of the new and revised items. Students were asked to rate each item, as if they were evaluating a course they had recently taken. They were asked to think aloud as they interpreted the items. Participants then indicated how they would answer using the 5-point scales. The researcher was particularly interested in items that students found difficult to understand or answer. Follow-up questions probed the potential causes for the difficulty. The interviews were audiotaped for subsequent analyses, and the researcher took notes during the interviews.

After the interviews, we analyzed the notes to discover consistent patterns in the ways that students understood and responded to the items. We made minor revisions to some items based on students' reactions.

Summary

Multiple information sources guided the revision process: the professional literature on teaching and learning; accreditation criteria from national and regional institutional accrediting agencies; IDEA staff; focus groups of IDEA users; expert panels comprised of psychometricians, statisticians, higher-education administrators, and experts in faculty development and evaluation; university faculty; and college students. From these resources suggestions were made to combine or eliminate some items and to create new ones.

Statistical Analyses

Description of the 2002 to 2011 Research Database

IDEA SRI data were retrieved from the years 2002-2011, which included 1,422,870 classes from 474 institutions. The *Diagnostic Form* (DF) was used in 929,521 classes and the *Short Form* (SF) in 493,254 classes. Several exclusions described in Table 1 were employed to produce a database consisting of 766,146 classes from 342 institutions (509,265 DF classes; 256,881 SF classes).[2]

Table 1

Historical Exclusions Employed in Creating the 2002-2011 Research Database

Exclusion	*n*	%
Classes with invalid number of students enrolled	20,373	1.4
Classes with fewer than 10 student responses	481,829	33.9
Classes from "novice" users of IDEA[a]	222,531	15.6
Classes with missing or invalid form type	95	< .001
Classes without an objective selected as "Essential" or "Important"	77,187	5.4
Classes without a matching institution record	12	< .001

Note. N = 1,422,870.
[a] Novice classes come from departments in their first year of using IDEA.

A final exclusion was then employed to limit the influence of any single institution to no more than 5% of the database. The 5% exclusion was run separately for DF and SF classes.[3] When classes from one institution are removed, the percent of the total for all other institutions increases. Consequently, multiple passes are required to continue making adjustments until no institution contributes more than 5% of the classes in the final dataset. For the DF two passes were required, resulting in 18,932 classes being dropped from two institutions. Three passes were required for the SF, resulting in 41,913 classes being dropped from five institutions. Once all exclusions were completed, 490,333 DF and 214,968 SF classes remained. Unless otherwise indicated, the analyses described in this report were performed on classes that used the DF.

Demographic Information about Participating Institutions

Demographic information about participating institutions is presented in Tables 2 and 3. Similar percentages of institutions came from Baccalaureate (21.7%) and Specialty Schools (22.2%) of art, music, and design. Slightly more were at the Master's level (27.7%), and the rest

[2] The total number of exclusions in Table 1 does not sum to total classes excluded because some classes were excluded for multiple reasons.
[3] All classes (N = 92,155) from the various Education Management Corporation (EDMC) institutions were considered as one institution for the 5% exclusion.

fell at the Associate (14.6%) and Doctoral (13.8%) levels. There was a greater representation of private (63.7%) than public (36.3%) institutions.

Table 2

Frequency and Percentage of Levels of Carnegie Classifications by Regional Location for Participating Institutions

	Associate		Baccalaureate		Masters		Doctoral		Special		Total	
Regional location	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Middle states	14	20.6	14	13.9	19	14.7	10	15.6	6	5.8	63	13.5
New England	0	0	5	5.0	5	3.9	4	6.3	1	1.0	15	3.2
North central	31	45.6	55	54.5	63	48.8	27	42.2	35	34.0	211	45.4
Northwest	2	2.9	4	4.0	5	3.9	2	3.1	7	6.8	20	4.3
Southern	18	26.5	22	21.8	27	20.9	14	21.9	39	37.9	120	25.8
Western	3	4.4	1	1.0	10	7.8	7	10.9	15	14.6	36	7.7
Total *n*	68		101		129		64		103		465	
Total %	14.6		21.7		27.7		13.8		22.2		100	

Table 3

Frequency and Percentage of Levels of Institutional Control by Regional Location for Participating Institutions

	Private		Public		Total	
Regional location	*n*	%	*n*	%	*n*	%
Middle states	41	13.9	22	13.0	63	13.5
New England	12	4.1	3	1.8	15	3.2
North central	133	44.9	78	46.2	211	45.4
Northwest	10	3.4	10	5.9	20	4.3
Southern	72	24.3	48	28.4	120	25.8
Western	28	9.5	8	4.7	36	7.7
Total *n*	296		169		465	
Total %	63.7		36.3		100	

Delivery Methods and Response Rates

Participating students completed either the DF or SF. Their data were included in the database only if their instructor completed an FIF. Although most student ratings were completed on paper, the percentage of online delivery increased across the years (see Table 4). Table 5 shows means and standard deviations for student response rates to surveys administered either on paper or online annually. The total response rate was higher for paper (81%) than online (66%).

Table 4

Frequency and Percentage of Paper and Online Survey Formats by Dataset Year

	Paper		Online	
Year	*n*	%	*n*	%
2002	19,745	99.3	134	0.7
2003	24,479	99.2	185	0.8
2004	28,878	99.2	230	0.8
2005	35,778	97.1	1,070	2.9
2006	37,676	94.2	2,302	5.8
2007	43,952	93.0	3,327	7.0
2008	55,023	91.5	5,100	8.5
2009	58,697	87.6	8,326	12.4
2010	68,733	83.7	13,377	16.3
2011	66,755	80.1	16,566	19.9
Total	439,716	89.7	50,617	10.3

Table 5

Means and Standard Deviations for Response Rates of Paper and Online Delivery by Dataset Year

	Paper			Online		
Year	*M*	*SD*	*n*	*M*	*SD*	*n*
2002	.81	.23	19,745	.67	.15	134
2003	.81	.36	24,479	.61	.16	185
2004	.80	.23	28,878	.65	.33	230
2005	.80	.25	35,778	.72	.16	1,070
2006	.80	.21	37,676	.71	.16	2,302
2007	.80	.25	43,952	.70	.17	3,327
2008	.80	.24	55,023	.68	.19	5,100
2009	.81	.33	58,697	.65	.19	8,326
2010	.81	.26	68,733	.65	.18	13,377
2011	.82	.21	66,755	.65	.18	16,566
Total	.81	.26	439,716	.66	.19	50,617

Student Responses to Individual Items

Table 6 presents means and standard deviations for the 47 items on the DF for both the 1998-2001 and 2002-2011 research datasets. Means for all items were higher in the latter time period, and most standard deviations decreased in magnitude. Average ratings have increased and become more restricted in range.

Table 6

Means and Standard Deviations for Student Ratings of Individual Items on the IDEA Diagnostic Form Across Two Time Periods

	1998-2001		2002-2011	
Item	*M*	*SD*	*M*	*SD*
Teaching methods				
1. Displayed personal interest in students	4.34	0.50	4.43	0.46
2. Helped students answer own questions	4.10	0.52	4.23	0.50
3. Scheduled work helpfully	4.20	0.48	4.31	0.46
4. Demonstrated importance of subject	4.32	0.45	4.41	0.44
5. Formed teams, discussion groups	3.52	1.03	3.74	0.91
6. Made clear how topics fit	4.20	0.51	4.32	0.48
7. Explained criticisms	3.78	0.57	4.01	0.55
8. Stimulated intellectual effort	3.86	0.57	4.06	0.54
9. Encouraged use of multiple resources	3.78	0.70	3.99	0.62
10. Explained clearly	4.13	0.61	4.24	0.58
11. Related to real life	4.22	0.58	4.33	0.54
12. Tests covered important points	4.28	0.49	4.36	0.46
13. Introduced stimulating ideas	4.03	0.58	4.19	0.54
14. Involved students in hands on activities	3.76	0.80	3.92	0.75
15. Inspired students to set high goals	3.76	0.62	3.97	0.59
16. Asked students to share experiences	3.69	0.79	3.89	0.73
17. Provided timely feedback	4.11	0.59	4.24	0.56
18. Asked students to help each other	3.79	0.64	3.98	0.59
19. Assessments required creativity	3.92	0.65	4.07	0.60
20. Encouraged student/faculty contact	3.90	0.63	4.09	0.57

Learning objectives				
21. Factual knowledge	3.94	0.52	4.13	0.47
22. Principles and theories	3.89	0.51	4.07	0.47
23. Applications	3.95	0.52	4.10	0.49
24. Professional skills, viewpoints	3.91	0.54	4.05	0.50
25. Team skills	3.45	0.82	3.61	0.74
26. Creative capacities	3.37	0.79	3.56	0.73
27. Broad liberal education	3.32	0.74	3.59	0.68
28. Communication skills	3.41	0.80	3.58	0.73
29. Find, use resources	3.58	0.60	3.78	0.56
30. Values development	3.44	0.69	3.67	0.65
31. Critical analysis	3.67	0.63	3.82	0.59
32. Interest in learning	3.74	0.56	3.88	0.54
Course ratings				
33. Amount of reading	3.20	0.74	3.25	0.71
34. Amount of other work	3.42	0.59	3.48	0.55
35. Difficulty of subject matter	3.42	0.58	3.47	0.55
Self-ratings				
36. Strong desire to take the course	3.66	0.67	3.71	0.66
37. Worked harder on this course than most	3.57	0.56	3.66	0.53
38. Wanted this instructor	3.40	0.67	3.54	0.67
39. Wanted course regardless of instructor	3.33	0.56	3.40	0.54
43. Usually work hard on academic work	3.64	0.31	3.81	0.33
Global ratings				
40. Increase positive attitude toward field	3.86	0.60	3.96	0.58
41. Excellent instructor	4.18	0.64	4.26	0.60
42. Excellent course	3.92	0.61	4.03	0.59
Additional method items				
44. Used variety of evaluation methods	3.83	0.60	3.95	0.55
45. Expected students to take responsibility	4.30	0.33	4.34	0.33
46. High achievement standards	4.13	0.41	4.19	0.40
47. Used educational technology	3.63	0.77	4.04	0.64

Characteristics of Courses

Instructors completed the *Faculty Information Form* (FIF) for each course section they taught. They provided information about their primary and secondary approaches to instruction, skills they required of students, circumstances they perceived to have either positively or negatively affected student learning, student level, and whether the course was team-taught and employed distance learning.

Primary and Secondary Approaches to Instruction

On the FIF, faculty identified which one of nine instructional methods represented their primary approach to the course. They also indicated their secondary approach. Response rates to these two questions were high, 98.9% and 97.2%, respectively. Table 7 presents frequency and percentage distributions of faculty responses.

The primary approach most frequently taken to instruction was lecture (51%), followed by discussion (11%). The reverse was true for secondary approaches. Primary instructional methods selected by fewer than 5% of instructors included seminar, laboratory, field experience, studio, multi-media, and practicum/clinic.

Table 7

Frequency and Percentage Distribution of Primary and Secondary Teaching Approaches

	Primary		Secondary	
Instructional approach	*n*	%	*n*	%
Lecture	249,451	50.9	71,408	14.6
Discussion/recitation	56,097	11.4	130,006	26.5
Seminar	23,322	4.8	14,217	2.9
Skill/activity	50,301	10.3	65,815	13.4
Laboratory	20,059	4.1	35,444	7.2
Field experience	3,137	0.6	8,830	1.8
Studio	12,170	2.5	4,220	0.9
Multi-media	5,463	1.1	14,802	3.0
Practicum/clinic	2,479	0.5	5,108	1.0
Other	62,666	12.8	126,884	25.9
Not rated	5,188	1.1	13,599	2.8

Course Requirements

Instructors were asked to report how much they required students to perform seven academic skills. They responded *None (or little) required, Some required,* or *Much required.*

Response rates ranged from 86.2% to 87.2%. Table 8 presents frequency and percentage distributions for faculty responses.

Over 50% of classes required at least some writing, oral communication, group work, and critical thinking. In contrast, fewer than 30% of classes required mathematical/quantitative and creative/artistic work. Surprisingly, only 44.1% required computer applications. Because of the substantial reliance upon computers in contemporary education, it remains unclear how instructors interpreted "computer applications" as a course requirement. Some instructors may interpret "computer applications" to mean requiring students to learn certain software packages pertinent to the subject matter (e.g., statistical packages for statistics courses, CAD software for mechanical engineering courses). On the other hand, some might consider general use of the computer, such as using word processing applications to write an essay, as "computer applications."

Table 8

Frequency and Percentage Distribution of Course Requirements Reported by Instructors

Response	Writing		Oral communication		Computer applications		Group work		Mathematical work		Critical thinking		Creative endeavor	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
None	87,420	17.8	111,165	22.7	207,354	42.2	155,807	31.8	282,071	57.5	44,798	9.1	288,057	58.7
Some	212,292	43.3	209,357	42.7	159,501	32.5	191,048	39.0	75,572	15.4	189,107	38.6	91,915	18.8
Much	127,700	26.0	105,189	21.5	56,887	11.6	78,365	16.0	66,124	13.5	192,327	39.2	42,756	8.7
Not rated	62,921	12.8	64,622	13.2	66,591	13.6	65,113	13.3	66,566	13.6	64,101	13.1	67,605	13.8

Note. Percentage columns (%) in tables do not always sum to 100 due to rounding.

Perceived Impact of Course Circumstances on Student Learning

Instructors rated the extent to which nine course circumstances had an impact on student learning. They responded *Had a positive impact on learning, Neither a positive nor a negative impact,* or *Had a negative impact on learning.* Response rates to these items ranged from 83.7% to 84.3%. Table 9 presents frequency and percentage distributions of faculty responses.

Over 50% of instructors believed previous experience teaching the course, desire to teach the course, and control over course management decisions had positive impacts on student learning. Faculty believed the highest negative impacts came from students' backgrounds and preparation (17.2%), physical facilities and/or equipment (11.8%), student enthusiasm for the course (11.4%), and student effort to learn (10.5%). In general, most instructors perceived course circumstances did not negatively impact learning.

Table 9

Frequency and Percentage Distribution of Perceived Impact of Course Circumstances on Student Learning

	Facilities		Experience		Changes to course		Desire to teach		Control		Student preparation		Student enthusiasm		Student effort		Support	
Response	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Positive impact	187,076	38.2	325,804	66.4	112,624	23.0	346,743	70.7	290,501	59.2	124,387	25.4	204,034	41.6	224,283	45.7	129,498	26.4
Neither	147,770	30.1	42,545	8.7	204,952	41.8	53,664	10.9	92,053	18.8	169,875	34.6	122,277	24.9	109,531	22.3	200,861	41.0
Negative impact	57,663	11.8	7,759	1.6	17,351	3.5	3,382	0.7	11,040	2.3	84,396	17.2	56,121	11.4	51,655	10.5	26,974	5.5
Can't judge	20,408	4.1	37,136	7.6	75,660	15.4	9,469	1.9	18,849	3.8	33,536	6.8	29,922	6.1	26,607	5.4	53,797	11.0
Not rated	77,416	15.8	77,089	15.7	79,746	16.3	77,075	15.7	77,890	15.9	78,139	15.9	77,979	15.9	78,257	16.0	79,203	16.2

Note. Percentage columns (%) do not always sum to 100 due to rounding. Facilities = Physical facilities and/or experiment; Experience = Your previous experience in teaching this course; Changes to course = Substantial changes in teaching approach, course assignments, content, etc.; Desire to teach = Your desire to teach this course; Control = Your control over course management decisions (objectives, texts, exams, etc.); Student preparation = Students' level of preparation for taking the course; Student enthusiasm = Students' level of enthusiasm for the course; Student effort = Students' level of effort to learn; Support = Technical/instructional support.

Principal Type of Student Enrolled in the Course

On the FIF, instructors described the principal type of student enrolled in the course. Options were (1) *First-year students/sophomores seeking to meet a "general education" or "distribution requirement";* (2) *First-year students/sophomores seeking to develop background needed for their intended specialization;* (3) *Upper level non-majors taking the course as a "general education" or "distribution requirement";* (4) *Upper level majors (in this or a related field of study) seeking competence or expertise in their academic/professional specialty;* (5) *Graduate or professional school students;* or (6) *Combination of two or more of the above types.* The response rate to this question was 84.6%. Table 10 presents frequency and percentage distributions for faculty responses to this item. Lower-level undergraduates comprised 41.1% of the students in the database, whereas 36.4% were upper level. Only 7.7% were primarily graduate/professional students.

Table 10

Frequency and Percentage Distributions of Principal Type of Student Enrolled in the Course

Principal type of student enrolled in course	*n*	%
Freshmen/sophomores meeting a general education requirement	123,801	25.2
Freshmen/sophomores developing background needed for intended major	78,080	15.9
Upper level non-majors	29,436	6.0
Upper level majors	99,974	20.4
Graduate or professional school students	37,958	7.7
Combination	45,591	9.3
Not rated	75,493	15.4

Team-taught Courses

Instructors were asked whether the course was team-taught. Response options were *Yes* and *No*. The response rate to this question was 84.5%. Only 5% identified the course as team-taught. Due to its irrelevance for the majority of instructors, this item might not be essential to include on the FIF.

Distance Learning

Instructors were also asked whether the course was taught through distance learning (*Yes* or *No*). The response rate to this question was 76.7% with only 1.1% responding "Yes." Given the ambiguity in the outdated term "distance learning," we concluded this item needed to be modified.

Summary

Nearly two-thirds of the institutions in the database were privately controlled, and nearly half were located in the North-Central U.S. All levels of Carnegie classifications and regions of the country were represented. Courses were taught mostly at the undergraduate level, and very few were team-taught or offered through distance learning. Most ratings were administered on paper, although the percentage of online surveys increased across the years. Response rates were higher for paper than online surveys, and mean student ratings were higher and more restricted in range than those in the 1998-2001 norming database.

Lecture remained the approach most frequently taken to instruction, followed by discussion. Very few instructors employed seminar, laboratory, field experience, studio, multi-media, and practicum/clinic. The majority of instructors required students to do at least some writing, oral communication, group work, and critical thinking. Most believed previous experience teaching the course, desire to teach the course, and control over course management decisions had positive impacts on student learning.

Validity

Validity refers to "the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests" (American Educational Research Association, 2014). In the context of student ratings, the "tests" are the survey instruments used to obtain student feedback about instruction. Interpretations and uses pertain to how ratings are applied in making summative decisions about teaching effectiveness and formative evaluations of how to improve teaching and the course. Too often student ratings are referred to as "student evaluations" and "course evaluations," but evaluations refer to judgments of worth based on multiple information sources. Student ratings are but one data source and should count no more than 30% to 50% of the overall teaching evaluation. Additional indicators of teaching quality should be considered. No single measure provides sufficient evidence to make a valid judgment about an instructor's overall effectiveness.

In this section, we present evidence that supports the revisions made to the IDEA SRI system. Several sources of validity evidence are provided: evidence based on relations to other variables, internal structure, test content, and expert judgments.

Validity Evidence Based on Relations of Student Ratings to Other Variables

Evidence of validity can be demonstrated in relationships between student ratings and external variables (Linn & Gronlund, 2000). *Concurrent validity* refers specifically to relations between a measure and some criterion assessed simultaneously. Faculty and students complete the FIF and DF, respectively, within the same semester, and so correlations between their ratings can provide evidence of a common construct.

Correlations Between Faculty Ratings of Importance and Student Ratings of Progress

Students rated their progress on each of the same 12 learning objectives their instructor rated for importance, using a five-point Likert scale (1 = *No apparent progress,* 2 = *Slight progress; I made small gains on this objective*, 3 = *Moderate progress; I made some gains on this objective,* 4 = *Substantial progress; I made large gains on this objective,* and 5 = *Exceptional progress; I made outstanding gains on this objective*). An indirect test of the validity of the IDEA SRI involves correlating students' mean progress for each objective with the instructor's ratings of importance. The highest correlations should be found in ratings of the same objectives if the following assumptions are valid (Hoyt, 1973, p. 376):

- Teaching was effective.
- Instructors paid careful attention to the identification of relevant objectives for each class.
- Student ratings of progress were valid.

Table 11 presents Pearson *r* correlations between faculty ratings of importance and student mean ratings of progress on each objective. The magnitudes and directions of the correlations are similar to those Hoyt and Lee (2002a) reported, and they show that students report more progress on objectives stressed by their instructor. The correlations between instructor and student ratings of the same objective (indicated in bold font) ranged from .07 to .36 (M = .20, SD = 0.08). The strongest correlations were found for Objectives 5, 6, and 8 (r = .31, .28, and .36, respectively), and the weakest was for Objective 12 (r = .07). The mean coefficient for off-diagonal (i.e., noncorresponding) correlations was close to zero (r = .03). These findings provide evidence for criterion-related validity in that students tend to report greater progress on objectives stressed by their instructor. The only exception is Objective 12, whose importance ratings had a weaker correlation with students' corresponding progress ratings than with students' ratings of four other objectives.

Table 11

Correlations Between Faculty Ratings of Importance and Student Ratings of Progress on 12 Learning Objectives

Item	FR1	FR2	FR3	FR4	FR5	FR6	FR7	FR8	FR9	FR10	FR11	FR12
SR1	**.16**	.09	.03	.10	-.01	-.03	-.04	-.10	-.04	-.00	-.07	.01
SR2	.12	**.13**	.07	.10	.00	-.02	-.06	-.10	-.04	.02	-.04	.02
SR3	-.01	.01	**.11**	.17	.06	.05	-.08	-.01	.02	.05	-.02	.03
SR4	.00	-.01	.08	**.21**	.07	.08	-.06	-.01	.02	.03	-.05	.02
SR5	-.12	-.09	.10	.16	**.31**	.06	-.04	.10	.08	.07	-.00	.04
SR6	-.24	-.20	-.02*	.11	.13	**.28**	.14	.27	.10	.09	.13	.06
SR7	-.12	-.13	-.08	.02	.07	.20	**.23**	.17	.03	.10	.11	.08
SR8	-.23	-.19	-.02	.05	.13	.14	.10	**.36**	.12	.13	.20	.08
SR9	-.06	-.06	.08	.11	.09	.04	-.04	.12	**.16**	.03	.07	.05
SR10	-.10	-.07	.05	.09	.11	.08	.02	.11	.08	**.21**	.12	.10
SR11	-.13	-.07	.03	.03	.06	.06	.02	.17	.08	.12	**.22**	.08
SR12	-.05	-.03	.05	.10	.07	.06	.00	.06	.05	.09	.07	**.07**

Note. SR = student ratings of progress. FR = faculty ratings of importance. Coefficients with absolute values $\geq .01$ are significant at $p < .001$.

Test-criterion relationships. In the context of IDEA SRIs, test-criterion validity evidence addresses whether student ratings of progress on relevant objectives predict how much students actually learned in the class. Benton, Duchon, and Pallett (2013) sought to answer that question by correlating students' progress ratings with their exam performance. Across multiple sections of the same course taught by the same instructor, they found student ratings of progress on relevant course objectives were positively correlated with exam scores whereas ratings on irrelevant objectives were not. Students who rated their progress as either exceptional or substantial generally outperformed those reporting moderate or less progress on course examinations.

Student Ratings of Progress on Relevant Objectives by Course Requirements

To further test relations of IDEA SRI to other variables, we replicated the analyses reported in Hoyt and Lee (2002a) by comparing student mean ratings of progress on relevant course objectives by the instructor's course requirements (see Table 12). As described by the previous authors:

> Specifically, if "writing" was emphasized, students should report above average progress on "Communication skills." If "critical thinking" was emphasized, above average progress should be reported on "Critical analysis." If "creative/artistic/design endeavor" was emphasized, students should report above average progress on "Creative capacities."

> And if "group work" was emphasized, student progress on "Team skills" should be relatively high (Hoyt & Lee, 2002a, p. 49).

In all cases, students reported greater progress on the relevant objective when instructors required *much* emphasis of the relevant skill rather than *none*. Strong effect sizes were found for communication skills (Cohen's d = .89 and .70, respectively) and creative capacities (Cohen's d = .89); a medium effect was found for critical thinking (Cohen's d = .61); and a small effect for team skills (Cohen's d = .22).

Table 12

Means and Standard Deviations for Student Ratings of Progress on Relevant Learning Objectives by Levels of Requirement for Various Academic Skills

Academic skill		*M*	*SD*	*n*	Cohen's *d*
Communication skills (Obj 8)	Course requirement: Writing				
	None	3.50	0.75	7,007	.89
	Some	3.75	0.64	78,207	
	Much	3.99	0.53	92,117	
	Course requirement: Oral communication				
		M	*SD*	*n*	Cohen's *d*
	None	3.62	0.67	18,692	.70
	Some	3.79	0.62	88,508	
	Much	4.02	0.54	69,666	
Creative capacities (Obj 6)	Course requirement: Creative endeavor				
		M	*SD*	*n*	Cohen's *d*
	None	3.71	0.65	19,384	.89
	Some	3.89	0.61	32,594	
	Much	4.22	0.52	32,941	
Team skills (Obj 5)	Course requirement: Group work				
		M	*SD*	*n*	Cohen's *d*
	None	3.63	0.72	106,532	.22
	Some	3.79	0.67	51,714	
	Much	3.79	0.69	16,992	
Critical thinking (Obj 11)	Course requirement: Critical thinking				
		M	*SD*	*n*	Cohen's *d*
	None	3.68	0.62	11,201	.61
	Some	3.87	0.55	70,278	
	Much	4.01	0.53	94,993	

Note. Cohen's *d* compares means between instructors placing "much" emphasis on the skill versus those placing "none."

Overall Measures by Course Circumstances

IDEA relies on three overall summary measures of teaching effectiveness. Progress on relevant objectives (PRO) is a weighted average of student mean ratings on relevant learning objectives identified by the instructor. Since objectives instructors identify as "essential" should receive greater teaching emphasis and account for more student progress, student ratings of *Essential* objectives are double weighted. Mean scores are also computed for ratings on two global items: (a) *Overall, I rate this instructor an excellent teacher*; and (b) *Overall, I rate this course as excellent*.

As further evidence for validity, we compared mean scores on the summary measures by instructor ratings of the impact of course circumstances (i.e., *positive impact, neither positive nor negative, negative impact)*, as was done in Hoyt and Lee (2002a). We hypothesized that mean student ratings would be higher for instructors who reported the course circumstance had a positive impact. Table 13 presents means and standard deviations by instructors' perceived impact of course circumstances on student learning. Cohen's *d* statistics are reported for differences between instructors perceiving a positive versus a negative impact on student learning. For the most part, instructors who rated course circumstances positively had higher ratings on all three overall measures than did those who rated them negatively. The strongest effects consistently occurred in ratings of excellence of the course.

Table 13

Means and Standard Deviations for Three Summary Measures by Instructors' Perceived Impact of Course Circumstances on Student Learning

	PRO			Excellence of teacher			Excellence of course		
Course circumstances	*M*	*SD*	*d*	*M*	*SD*	*d*	*M*	*SD*	*d*
Physical facilities and/or equipment									
Positive (*n* = 187,019)	53.38	8.20	.09	51.69	9.18	.13	52.45	9.51	.16
In between (*n* = 147,737)	53.00	8.11		51.23	9.21		51.44	9.62	
Negative (*n* = 57,652)	52.62	8.38		50.52	9.75		50.95	10.05	
Previous experience in teaching this course									
Positive (*n* = 325,722)	53.47	7.95	.39	51.79	8.92	.45	52.36	9.35	.52
In between (*n* = 42,535)	52.10	8.80		49.64	10.18		50.14	10.29	
Negative (*n* = 7,756)	50.35	9.63		47.78	11.26		47.50	11.29	
Substantial changes in teaching approach, course assignments, content, etc.									
Positive (*n* = 112,597)	53.25	8.24	.15	51.47	9.31	.19	52.17	9.66	.22
In between (*n* = 204,891)	53.41	7.99		51.67	9.01		52.17	9.43	
Negative (*n* = 17,345)	51.97	8.84		49.64	10.22		49.99	10.47	
Desire to teach this course									
Positive (*n* = 346,650)	53.40	8.06	.32	51.72	9.06	.38	52.33	9.44	.50
In between (*n* = 53,653)	51.51	8.68		49.08	10.08		49.11	10.15	
Negative (*n* = 3,378)	50.83	9.21		48.24	10.73		47.59	11.04	
Control over course management decisions (objectives, texts, exams, etc.)									
Positive (*n* = 290,419)	53.54	7.97	.36	51.85	8.95	.37	52.56	9.34	.56
In between (*n* = 92,035)	52.27	8.54		50.24	9.78		50.45	9.96	
Negative (*n* = 11,036)	50.68	9.26		48.48	10.85		47.25	11.11	
Students' level of preparation for taking the course									
Positive (*n* = 124,328)	54.09	8.09	.37	52.24	9.05	.35	53.34	9.26	.50
In between (*n* = 169,840)	53.49	7.84		51.91	8.82		52.46	9.29	
Negative (*n* = 84,385)	51.03	8.71		48.90	10.14		48.49	10.18	
Students' level of enthusiasm for the course									
Positive (*n* = 203,961)	54.33	7.78	.56	52.75	8.64	.56	54.05	8.92	.86
In between (*n* = 122,261)	52.75	8.08		50.85	9.21		51.05	9.30	
Negative (*n* = 56,111)	49.82	8.87		47.64	10.39		46.10	10.25	
Students' level of effort to learn									
Positive (*n* = 224,209)	54.01	7.95	.48	52.37	8.88	.50	53.35	9.22	.67
In between (*n* = 109,514)	52.88	8.01		51.07	9.08		51.26	9.39	
Negative (*n* = 51,647)	50.09	8.86		47.76	10.35		47.06	10.22	
Technical/instructional support									
Positive (*n* = 129,451)	53.00	8.32	.02	51.25	9.40	.07	52.00	9.67	.07
In between (*n* = 200,812)	53.31	8.05		51.56	9.10		51.95	9.55	
Negative (*n* = 26,968)	52.87	8.45		50.56	9.83		51.34	10.02	

Note. PRO = Progress on relevant objectives. Cohen's *d* compares means between instructors who reported the course circumstance had a positive effect on student learning versus those who reported a negative effect.

Student Motivation by Student Types

As was done in Hoyt and Lee (2002a), we compared student motivation to take the course by student types. Two items on the DF were used to measure students' motivation: *I had a strong desire to take this course* (item 36), and *I really wanted to take this course regardless of who taught it* (item 39). If instructors correctly identify and report the principal type of student enrolled in the course and students accurately disclose their motivation, one would expect differences on the two motivation items. That is, students taking the course to gain background or expertise in their intended area of specialization should be more motivated than those taking it to fulfill a general education or distribution requirement. Table 14 presents descriptive statistics for the two measures of motivation. Ratings of the *desire to take the course* were higher (Cohen's $d = .77$) for lower- and upper-division students specializing in the subject matter ($M = 3.89$, $SD = 0.60$) than for lower- and upper-division students fulfilling general education requirements ($M = 3.64$, $SD = 0.64$). Similarly, ratings of the *desire to take the course regardless of who taught it* were higher (Cohen's $d = .62$) for students in the major ($M = 3.51$, $SD = 0.51$) than for general education students ($M = 3.19$, $SD = 0.53$).

Table 14

Means and Standard Deviations for Student Motivation Variables by Principal Type of Student

Type of student	Item 36			Item 39		
	M	*SD*	*n*	*M*	*SD*	*n*
Lower division, general education	3.38	0.64	123,800	3.17	0.53	123,800
Upper division, general education	3.57	0.66	29,436	3.25	0.53	29,436
Lower division, specialized	3.88	0.61	78,078	3.53	0.52	78,078
Upper division, specialized	3.90	0.59	99,972	3.50	0.50	99,972
Graduate/professional	3.92	0.57	37,930	3.52	0.48	37,930

Note. Item 36 = I had a strong desire to take this course. Item 39 = I really wanted to take this course regardless of who taught it.

Summary

Correlations between faculty and student ratings of learning objectives are highest for ratings of the same objectives. Students report greater progress on relevant learning objectives when the instructor requires much emphasis of a relevant skill rather than none. In general, instructors who rate course circumstances positively have higher ratings on overall summary measures than do those who rate them negatively. Students who take a course to gain background or expertise in their area of specialization tend to be more motivated than those taking it to fulfill a general education or distribution requirement.

Validity Evidence Based on Internal Structure

Evidence for validity can also be shown in an instrument's internal structure and the relationship among items. One question concerns whether we can confirm Hoyt and Lee's (2002a) finding that faculty and student ratings are each multi-dimensional. An assumption of the IDEA system is that instructors are best qualified to judge the relevance of learning objectives for their course. If we could confirm that faculty ratings are multi-dimensional, then it would demonstrate instructors can discern the relevance of the objectives for their course. It is also assumed that students are capable of distinguishing how much progress they made on the 12 learning objectives and how frequently the instructor applied each of the 20 teaching methods. Confirming that student ratings are multi-dimensional would give support to those assumptions.

Correlations Among Faculty Ratings of Learning Objectives

Table 15 presents correlations among faculty ratings of the 12 learning objectives. The general pattern was similar to that reported in Hoyt and Lee (2002a). The highest correlations were found between Objective 12 and Objectives 9-11. Instructors who selected "interest in learning" also tended to emphasize finding resources ($r = .52$), values development ($r = .51$), and critical thinking ($r = .46$). Moderately positive correlations were also observed between Objectives 1 and 2 ($r = .41$); 6 and 7 ($r = .41$); and 8, 9, and 11 ($r = .41$, .42, and .45, respectively).

The moderately high correlations between Objectives 1 and 2 and between Objectives 9 and 12 make sense. Objectives 1 and 2 were developed to assess student progress on cognitive learning outcomes (Hoyt & Cashin, 1977); Objectives 9 and 12 focus on lifelong learning (Hoyt et al., 1999). Among faculty members who selected Objective 1 as *Important*, 76.6% also identified Objective 2 as either *Important* or *Essential*. Among those who selected Objective 1 as *Essential*, 85.7% rated Objective 2 as either *Important* or *Essential*. The percentages were not as high when cross-tabulating faculty responses to Objectives 9 and 12; nonetheless, there was some overlap. Three in five faculty members who identified Objective 9 as *Important* selected Objective 12 as either *Important* or *Essential*. Three quarters of instructors selecting Objective 9 as *Essentia*l rated Objective 12 as either *Important* or *Essential*.

Table 15

Pearson r Correlations Between Faculty Ratings of Relevance on 12 Learning Objectives

Item	FR1	FR2	FR3	FR4	FR5	FR6	FR7	FR8	FR9	FR10	FR11	FR12
FR1	—											
FR2	.41	—										
FR3	.10	.25	—									
FR4	.10	.11	.29	—								
FR5	.01	.07	.23	.30	—							
FR6	-.07	-.01	.11	.27	.31	—						
FR7	.02	.01	-.01	.07	.22	.41	—					
FR8	-.19	-.13	.05	.07	.29	.32	.28	—				
FR9	.03	.07	.28	.25	.34	.27	.22	.41	—			
FR10	.04	.12	.21	.20	.36	.30	.33	.29	.37	—		
FR11	-.12	.03	.17	.04	.20	.21	.27	.45	.42	.38	—	
FR12	.13	.19	.29	.25	.37	.33	.36	.34	.52	.51	.46	—

Note. FR = faculty ratings of importance. All coefficients are significant at $p < .001$.

Factor Structure of Faculty Ratings of Learning Objectives

We conducted principal components analysis with varimax rotation to confirm the multi-dimensionality of faculty ratings on the 12 learning objectives. Hoyt and Lee (2002a) reported three underlying factors comprised of *Intellectual Development* (Objectives 11, 12, 10, 7, and 8), *Professional Preparation* (Objectives 4, 5, and 3), and *Basic Cognitive Development* (Objectives 1 and 2). Objectives 6 and 9 loaded on both *Intellectual Development* and *Professional Preparation*.

In the current analysis, factors with eigenvalues greater than 1.0 were extracted and rotated. The scree plot showed a large drop off from the first to the second factor and gently declined to the third and fourth factors. After that point, the plot leveled off. Table 16 presents rotated component matrix coefficients.

The first rotated factor explained 22% of the variance (eigenvalue = 2.64) and was comprised of objectives pertaining to *Intellectual Development*, including critical thinking (Objective 11), information literacy (Objective 9), lifelong learning (Objective 12), communication skills (Objective 8), and values development (Objective 10). The second factor explained 14.33% (eigenvalue = 1.79) of the variance and could be described as *Professional Development*, specifically professional skills development (Objective 4), application of course material (Objective 3), and team skills (Objective 5). The third factor, which explained 13.22% (eigenvalue = 1.59) of the variance, concerned *Basic Cognitive Background*--factual knowledge (Objective 1) and principles and theories (Objective 2). Finally, the fourth factor explained an additional 12.96% (eigenvalue = 1.56) of the variance and dealt with *Cultural/Creative Development*, which included Objectives 7 (cultural development) and 6 (creative/artistic

capacities). The current findings confirm that instructors are able to make distinctions about the relevance of the learning objectives for their courses.

Table 16

Rotated Factor Loadings for Faculty Ratings of Importance on Learning Objectives

	Factor loading			
Learning objective	1	2	3	4
11. Critical analysis	**.82**	-.08	-.08	.05
9. Find, use resources	**.68**	.33	.03	.05
12. Interest in learning	**.68**	.25	.25	.27
8. Communication skills	**.64**	.04	-.34	.25
10. Values development	**.56**	.22	.16	.32
4. Professional skills/viewpoints	-.03	**.82**	.05	.12
3. Applications	.33	**.59**	.23	-.34
5. Team skills	.29	**.57**	-.03	.28
1. Factual knowledge	-.10	.02	**.82**	.07
2. Principles/theories	.10	.09	**.79**	-.07
7. Broad liberal education	.27	-.06	.08	**.79**
6. Creative capacities	.17	.37	-.12	**.68**

Note. Boldface indicates highest factor loadings.

Correlations Among Student Ratings of Teaching Methods

Correlations among student ratings of teaching methods can be found in Appendix A. All the teaching methods were positively correlated with one another, with *r* ranging from .28 to .91 (M = .68).

Factor Structure of Student Ratings of Teaching Methods

Hoyt and Lee (2002a) reported a two-factor structure comprised of teaching methods related to the instructor's role in *Knowledge Transmission* and the student's role in *Knowledge Acquisition*. To confirm this structure, we conducted principal components analysis with varimax rotation on student ratings of how frequently their instructor used each of the 20 teaching methods, ranging from 1 (*hardly ever*) to 5 (*almost always*). Rotated component matrix coefficients are presented in Table 17. The rotated two-factor solution was similar to what Hoyt and Lee (2002a) reported. In the current analysis, 14 teaching methods loaded on the first factor (eigenvalue = 9.45), which explained 47.27% of the variance and focused on the instructor's role in *Knowledge Transmission.* The second factor (eigenvalue = 6.18), which explained 30.89% of the variance, emphasized the student's role in what we preferred to call *Knowledge*

Construction. The multi-dimensionality of the ratings indicated that students do discriminate broadly among the teaching methods when they rate how frequently each method occurs in the classroom.

Table 17

Rotated Factor Loadings for Student Ratings of Teaching Methods

	Factor loading	
Teaching method item	1	2
TM 10	**.88**	.28
TM 4	**.85**	.36
TM 6	**.85**	.38
TM 12	**.84**	.16
TM 2	**.81**	.46
TM 1	**.79**	.43
TM 17	**.79**	.20
TM 3	**.79**	.36
TM 13	**.79**	.49
TM 8	**.75**	.48
TM 7	**.71**	.54
TM 20	**.69**	.45
TM 11	**.68**	.43
TM 15	**.67**	.64
TM 5	.10	**.86**
TM 14	.28	**.84**
TM 16	.41	**.78**
TM 19	.47	**.75**
TM 18	.51	**.74**
TM 9	.43	**.71**

Note. Boldface indicates higher factor loadings.

Correlations Among Student Ratings of Learning Objectives

Correlations among student ratings of learning objectives are presented in Appendix A. Student-reported progress on all the learning objectives were positively correlated with one another, with *r* ranging from .45 to .91 ($M = .70$). The correlation between Objectives 1 and 2 and that between Objectives 3 and 4 were the highest. Students who reported to have made progress on one of the two cognitive learning outcomes (Objectives 1 and 2) tended to do well on the other ($r = .91$). Those who learned about application of course material (Objective 3) also made progress on developing professional skills and perspectives (Objective 4) ($r = .91$).

Factor Structure of Student Ratings of Learning Objectives

The principal components analysis we performed on student ratings of progress on the 12 learning objectives essentially replicated that of Hoyt and Lee (2002a). Following varimax rotation, two factors emerged as shown by the component matrix coefficients in Table 18. The first factor (eigenvalue = 5.46), which explained 45.48% of the variance, seemed to focus primarily on *Expressiveness and Intellectual Development* by virtue of the high loadings from ratings on communication skills (Objective 8), creative/artistic capacities (Objective 6), cultural activities (Objective 7), values development (Objective 10), and critical thinking (Objective 11). The second factor (eigenvalue = 4.6) explained 38.33% of the variance and concerned *Basic Cognitive Background and Application of Learning*, with high loadings from ratings on factual knowledge (Objective 1), principles and theories (Objective 2), application (Objective 3), and professional skill development (Objective 4). As was the case with Hoyt and Lee's analysis, Objective 12 had nearly equivalent loadings on both factors. The results show that students can differentiate the progress made on broad categories of achievement.

Table 18

Rotated Factor Loadings for Student Ratings of Progress on Learning Objectives

	Factor loading	
Learning objective	1	2
8. Communication skills	**.91**	.24
6. Creative capacities	**.90**	.23
7. Broad liberal education	**.81**	.34
10. Values development	**.79**	.47
11. Critical analysis	**.76**	.50
5. Team skills	**.71**	.32
9. Find, use resources	**.70**	.55
12. Interest in learning	.67	.66
1. Factual knowledge	.24	**.92**
2. Principles/theories	.29	**.92**
3. Applications	.44	**.84**
4. Professional skills/viewpoints	.46	**.82**

Note. Boldface indicates higher factor loadings.

Relationships Between Teaching Methods and Learning Objectives

Another assumption about the IDEA system's internal structure is that the correlations between teaching methods and learning objectives are generally distinctive for each objective

(Hoyt & Lee, 2002a). The only exceptions are for Objectives 1 and 2 (basic cognitive background) and Objectives 3 and 4 (applications, professional skills and viewpoints), which tend to have identical lists of related teaching methods due to their conceptual similarity.

To investigate which teaching methods are most important for explaining student progress reported on each learning objective, we employed Bayesian Model Averaging (BMA). BMA is an ensemble technique that tests multiple models to obtain better predictive performance than could be obtained with a single model (Hoeting, Madigan, Raftery, & Volinsky, 1999). It provides estimated probabilities that the frequency of each teaching method is associated with a given learning objective. A criterion (Schwartz Bayesian Criterion) is used for model selection among the finite set of models (2 to the kth power, where k = the number of explanatory variables). The SBC introduces a penalty term for increasing the number of predictors. We selected the best 100 models, based on the SBC criterion. Only classes where the instructor rated the learning objective as relevant were included in the analysis. Separate analyses were conducted on each learning objective.

Tables of estimated probabilities and regression parameters (weighted coefficients), broken out by class size, are presented for each learning objective in Appendix B. Table 19 summarizes the significant explanatory variables (indicated by item number) included in the "best" full models for each objective by class size. The following four teaching methods were included in no models, which calls into question their validity for diagnostic feedback about learning objectives:

1. Displayed a personal interest in students and their learning

3. Scheduled course work (class activities, tests, projects) in ways which encouraged students to stay up to date in their work

17. Provided timely and frequent feedback on tests, reports, projects, etc. to help students improve

20. Encouraged faculty-student interaction outside of class (office visits, phone calls, e-mail, etc.)

Table 19

Relevant Teaching Methods for Progress on Learning Objectives by Class Size

Learning objective	Class size			
	Small (10-15)	Medium (15-34)	Large (35-49)	Very large (50+)
1. Gaining factual knowledge (terminology, classifications, methods, trends)	8 (6, 10, 12, 13)	6, 8 (12, 13)	6, 8 (12, 13)	6, 8 (12, 13, 15)
2. Learning fundamental principles, generalizations, or theories	8 (6, 12, 13)	8 (6, 12, 13)	8 (6, 12, 13)	8 (6, 11, 12, 18)
3. Learning to apply course material (to improve thinking, problem solving, and decisions)	(4, 8, 11, 14, 15)	(2, 4, 8, 11, 14, 15, 18)	15 (2, 4, 6, 8, 11, 18)	11, 15, 18 (2, 4, 6, 8)
4. Developing specific skills, competencies, and points of view needed by professionals in the field most closely related to this course	15 (4, 6, 14)	15 (4, 6, 8, 14, 18)	6, 15 (4, 8, 11, 14, 18)	15 (4, 6, 8, 11, 14, 18)
5. Acquiring skills in working with others as a member of a team	5, 14, 15, 18 (8, 10)	5, 14, 15, 18 (10)	5, 14, 15, 18	5, 14, 15 (18)
6. Developing creative capacities (writing, inventing, designing, performing in art, music, drama, etc.)	7, 15, 19 (10, 13)	7, 15, 19 (10, 13)	7, 10, 13,15, 19 (14, 16)	7, 10, 13, 15, 16, 19
7. Gaining a broader understanding and appreciation of intellectual/cultural activity (music, science, literature, etc.)	7, 13, 15, 16 (6, 10, 19)	13, 16, 19 (6, 7, 10, 15)	7, 10, 13, 15, 16, 19 (6)	7, 10, 13, 16 (15, 19)
8. Developing skill in expressing myself orally or in writing	7, 16, 19 (5, 9, 10, 15)	7, 16, 19 (5, 9, 10, 15)	7, 15, 16, 19 (8, 9, 10)	7, 15, 16, 19 (10, 13)
9. Learning how to find and use resources for answering questions or solving problems	9 (2, 8, 10, 15, 18)	9 (2, 8, 15, 18)	9, 15 (2, 8, 18)	9, 15 (8, 18)
10. Developing a clearer understanding of, and commitment to, personal values	15, 16 (8, 11, 13)	15, 16 (8, 11, 13)	4, 13, 15, 16 (19)	4, 15, 16 (19)
11. Learning to analyze and critically evaluate ideas, arguments, and points of view	8, 13, 16 (2, 9, 19)	8, 16, 19 (2, 9, 13)	2, 8, 16, 19 (13)	8, 13, 16, 19 (2, 4)
12. Acquiring an interest in learning more by asking my own questions and seeking answers	8 (2, 13, 15, 16, 18)	8, 18 (2, 13, 15, 16)	2, 8, 13, 15, 16 (18)	2, 8, 13, 15 (16, 18)

Note. Item numbers within parentheses had standardized regression coefficients $\geq$ .05 and < .10. Those outside parentheses had coefficients $\geq$ .10.

Teaching Methods	
1. Displayed personal interest in students and their learning	12. Gave tests, projects, etc. that covered the most important points of the course
2. Found ways to help students answer their own questions	13. Introduced stimulating ideas about the subject
3. Scheduled course work (class activities, tests, projects) in ways which encouraged students to stay up-to-date in their work	14. Involved students in "hands on" projects such as research, case studies, or "real life" activities
4. Demonstrated the importance and significance of the subject matter	15. Inspired students to set and achieve goals which really challenged them
5. Formed "teams" or "discussion groups" to facilitate learning	16. Asked students to share ideas and experiences with others whose backgrounds and viewpoints differ from their own
6. Made it clear how each topic fit into the course	17. Provided timely and frequent feedback on tests, reports, projects, etc. to help students learn
7. Explained the reasons for criticisms of students' academic performance	18. Asked students to help each other understand ideas or concepts
8. Stimulated students to intellectual effort beyond that required by most courses	19. Gave projects, tests, or assignments that required original or creative thinking
9. Encouraged students to use multiple resources (e.g., data banks, library holdings, outside experts) to improve understanding	20. Encouraged student-faculty interaction outside of class (office visits, phone calls, email, etc.)
10. Explained course material clearly and concisely	
11. Related course material to real life situations	

Relationships Between Teaching Methods and Overall Summary Measures

As further evidence for internal structure, we applied BMA to examine which teaching methods are most important for explaining variance in student ratings of the two summary measures: *Overall, I rate this instructor an excellent teacher* and *Overall, I rate this course as excellent*. Students responded to these two items using the scale, 1 = *Definitely False*, 2 = *More False than True*, 3 = *In Between*, 4 = *More True than False*, and 5 = *Definitely True*. Tables of estimated probabilities and regression parameters, broken down by class size, appear in Appendix C.

Seven teaching methods were significantly related to either one or both of the overall summary measures. Table 20 shows the seven teaching methods associated with each outcome, categorized by four major areas: organization, clarity, enthusiasm/expression, and rapport/interactions. Across multiple factor-analytic studies of student ratings (Braskamp & Ory, 1994; Feldman, 1989; Hativa, Barak, & Simhi, 2001; Marsh, 1987; Murray, 1997), those four broad teacher behaviors are most highly correlated with ratings of teaching effectiveness (as summarized in Hativa, 2013a). The seven teaching methods in Table 20 were considered most important, then, for making improvements in ratings of the teacher and the course.

Table 20

Teaching Methods Related to Overall Summary Measures of Course and Instructor

	Overall summary measure	
Teaching method category	Excellence of course	Excellence of instructor
Organization	6. Made it clear how each topic fit into the course	
Clarity	10. Explained course material clearly and concisely	10. Explained course material clearly and concisely
Enthusiasm/expression	13. Introduced stimulating ideas about the subject	13. Introduced stimulating ideas about the subject
	15. Inspired students to set and achieve goals which really challenged them 4. Demonstrated the importance and significance of the subject matter	
Rapport/interactions		1. Displayed personal interest in students and their learning 2. Found ways to help students answer their own questions

Summary

The current findings confirm that faculty and student ratings of learning objectives and student ratings of teaching methods are each multi-dimensional. Moreover, the correlations between teaching methods and learning objectives are generally distinctive for each objective. Seven teaching methods are specifically essential for explaining variance in overall ratings of the teacher and course. Three teaching methods are not strongly associated with learning objectives or overall measures: Items 3 (scheduled coursework), 17 (provided timely and frequent feedback), and 20 (encouraged student-faculty interaction outside of class).

Evidence Based on Test Content

Evidence based on test content concerns the development of the wording and format for items on an instrument. The initial development of the existing 47 DF items has been described previously (Hoyt, 1973a, 1973b; Hoyt & Cashin, 1977; Hoyt & Lee, 2002a). Of the 13 learning objectives contained in this revision (*Diagnostic Feedback*), six first appeared in similar form in *Student Reactions to Instructors and Courses* (1969). One new objective was added to *Student Reactions to Instruction and Courses* (2nd ed., 1972). The number and wording of learning objectives remained the same in *IDEA Survey Form—Student Reactions to Instruction and Courses* (1975, 1988). Then two more appeared in *The IDEA System for Obtaining Student Ratings of Instructors and Courses* (1998). Table 21 shows the progression in wording of 13 objectives across time.

Table 21

Progression in Wording of IDEA Learning Objectives Across Time

1969	1972/1975/1988	1998	Diagnostic Feedback
▪ Gaining factual knowledge (terminology, classifications, methods, trends)	▪ Gaining factual knowledge (terminology, classifications, methods, trends)	▪ Gaining factual knowledge (terminology, classifications, methods, trends)	▪ Gaining a basic understanding of the subject (e.g., factual knowledge, methods, principles, generalizations, theories)
▪ Learning fundamental principles, generalizations, or theories	▪ Learning fundamental principles, generalizations, or theories	▪ Learning fundamental principles, generalizations, or theories	
▪ Learning to apply principles to solve practical problems	▪ Learning to apply course material to improve rational thinking, problem-solving, and decision making	▪ Learning to *apply* course material (to improve thinking, problem-solving, and decisions)	▪ Learning to *apply* course material (to improve thinking, problem-solving, and decisions)
▪ Learning attitudes and behavior characteristics of professionals in the field most closely related to this course	▪ Developing specific skills, competencies and points of view needed by professionals in the field most closely related to this course	▪ Developing specific skills, competencies and points of view needed by professionals in the field most closely related to this course	▪ Developing specific skills, competencies and points of view needed by professionals in the field most closely related to this course
▪ Developing skill in effective communication	▪ Developing skill in expressing myself orally or in writing	▪ Developing skill in expressing myself orally or in writing	▪ Developing skill in expressing myself orally or in writing
▪ Gaining a broader understanding and appreciation of intellectual-cultural matters (music, science, literature, etc.)	▪ Gaining a broader understanding and appreciation of intellectual-cultural activity (music, science, literature, etc.)	▪ Gaining a broader understanding and appreciation of intellectual/cultural activity (music, science, literature, etc.)	▪ Gaining a broader understanding and appreciation of intellectual/cultural activity (music, science, literature, etc.)
	▪ Developing creative capacities	▪ Developing creative capacities (writing, inventing, designing, performing in art, music, drama, etc.)	▪ Developing creative capacities (inventing; designing; writing; performing in art, music, drama, etc.)
		▪ Acquiring skills in working with others as a member of a team	▪ Acquiring skills in working with others as a member of a team
		▪ Learning how to find and use resources for answering questions or solving problems	▪ Learning how to find, evaluate, and use resources to explore a topic in depth
		▪ Learning to *analyze* and *critically evaluate* ideas, arguments, and points of view	▪ Learning to *analyze* and *critically evaluate* ideas, arguments, and points of view

	▪ Developing knowledge and understanding of diverse perspectives, global awareness, or other cultures ▪ Developing ethical reasoning and/or ethical decision making ▪ Learning to apply knowledge and skills to benefit others or serve the public good ▪ Learning appropriate methods for collecting, analyzing, and interpreting numerical information

Of the 19 teaching methods in this revision, six first appeared in 1969 and one was added to the forms produced in 1972, 1975, and 1988. The 1998 version included nine new teaching methods. Table 22 shows the changes in wording of items across time.

Table 22

Progression in Wording of IDEA Teaching Methods Across Time

1969	1972/1975/1988	1998	Diagnostic Feedback
▪ He explained course material clearly, and explanations were to the point	▪ Explained course material clearly, and explanations were to the point	▪ Explained course material clearly and concisely	▪ Explained course material clearly and concisely
▪ He stimulated students to intellectual effort beyond that required by most courses	▪ Stimulated students to intellectual effort beyond that required by most courses	▪ Stimulated students to intellectual effort beyond that required by most courses	▪ Stimulated students to intellectual effort beyond that required by most courses
▪ He introduced stimulating ideas about the subject	▪ Introduced stimulating ideas about the subject	▪ Introduced stimulating ideas about the subject	▪ Introduced stimulating ideas about the subject
▪ He related course material to real life situations	▪ Related course material to real life situations	▪ Related course material to real life situations	▪ Related course material to real life situations
▪ He explained the reasons for his criticism of students' academic performance	▪ Explained the reasons for his criticisms of students' academic performance ('72, '75)	▪ Explained the reasons for criticisms of students' academic performance	▪ Provided meaningful feedback on students' academic performance
	▪ Explained the reasons for criticisms of students' academic performance ('88)		
	▪ Displayed a personal interest in me and my learning	▪ Displayed a personal interest in students and their learning	▪ Displayed a personal interest in students and their learning
	▪ Found ways to help students answer their own questions	▪ Found ways to help students answer their own questions	▪ Found ways to help students answer their own questions
	▪ Demonstrated the importance and significance of his subject matter	▪ Demonstrated the importance and significance of the subject matter	▪ Demonstrated the importance and significance of the subject matter
		▪ Formed "teams" or "discussion groups" to facilitate learning	▪ Formed teams or groups to facilitate learning
		▪ Encouraged students to use multiple resources (e.g., data banks, library holdings, outside experts) to improve understanding	▪ Encouraged students to use multiple resources (e.g., Internet, library holdings, outside experts) to improve understanding
		▪ Involved students in "hands on" projects such as research, case studies, or "real life" activities	▪ Involved students in hands-on projects such as research, case studies, or real life activities

	▪ Inspired students to set and achieve goals which really challenged them ▪ Asked students to share ideas and experiences with others whose backgrounds and viewpoints differ from their own ▪ Asked students to help each other understand ideas or concepts ▪ Gave projects, tests, or assignments that required original or creative thinking	▪ Inspired students to set and achieve goals which really challenged them ▪ Asked students to share ideas and experiences with others whose backgrounds and viewpoints differ from their own ▪ Asked students to help each other understand ideas or concepts ▪ Gave projects, tests, or assignments that required original or creative thinking ▪ Created opportunities for students to apply course content outside the classroom ▪ Helped students to interpret subject matter from diverse perspectives (e.g., different cultures, religions, genders, political views, etc.) ▪ Encouraged students to reflect on and evaluate what they have learned

Course characteristics did not appear until the 1972 edition; student characteristics were first introduced in 1998. Table 23 shows the progression of these items.

Table 23

Progression in Wording of IDEA Student and Course Characteristics Across Time

1972/1973/1988	1998	Diagnostic Feedback
▪ Amount of reading	▪ Amount of reading	▪ Amount of coursework
▪ Amount of work in other (non-reading) assignments	▪ Amount of work in other (non-reading) assignments	
▪ Difficulty of subject matter	▪ Difficulty of subject matter	▪ Difficulty of subject matter
	▪ I really wanted to take this course regardless of who taught it.	▪ I really wanted to take this course regardless of who taught it.
	▪ As a rule, I put forth more effort than other students on academic work.	▪ As a rule, I put forth more effort than other students on academic work.
	▪ My background prepared me well for this course's requirements. (Short Form only)	▪ My background prepared me well for this course's requirements.
		▪ When this course began I believed I could master its content.

Two overall summary measures were added to the 1988 version that have been retained through the current revision: *Overall, I rate this instructor an excellent teacher*, and *Overall, I rate this an excellent course*.

Validity Evidence Based on Expert Judgments

Validity evidence can also be found in the judgments of experts (American Educational Research Association et al., 1999). We considered several sources of expertise: faculty ratings of the importance of IDEA learning objectives and input from expert panels and focus groups.

Faculty Ratings of the Importance of Learning Objectives

Instructors provided expertise about the importance of each of the IDEA learning objectives when they rated them for relevance to their course, using the scale *Minor or No Importance, Important*, or *Essential*. Table 24 presents descriptive statistics for faculty ratings of the objectives.

Table 24

Descriptive Statistics for Faculty Ratings of the Importance of Learning Objectives

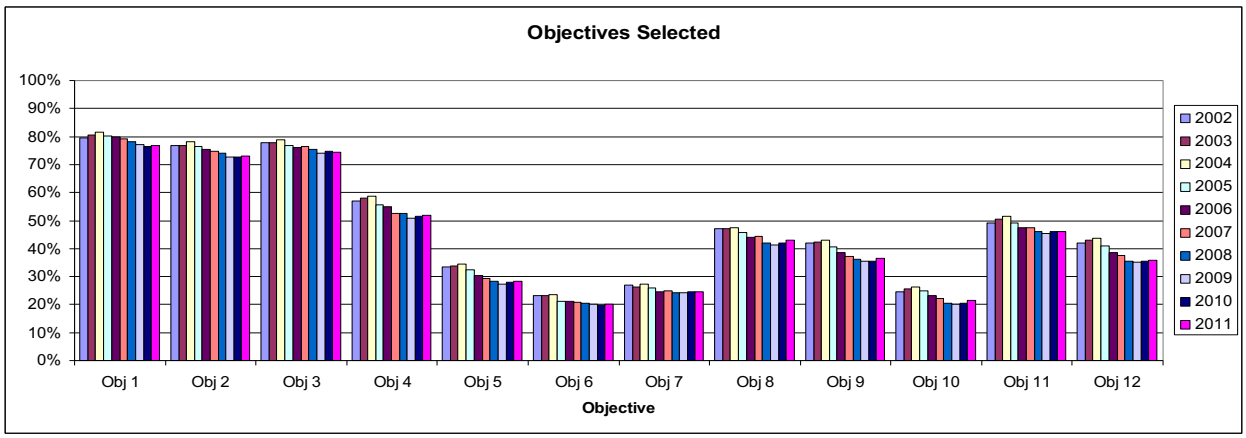
Learning objectives	% Important	% Essential	% Essential or Important	*M*	*SD*	*N*
1. Factual knowledge	28.6	47.9	76.5	2.22	0.86	481,929
2. Principles/theories	32.0	40.4	72.4	2.10	0.86	480,097
3. Applications	36.9	36.9	73.8	2.08	0.84	480,191
4. Professional skills/viewpoints	27.4	23.5	50.9	1.70	0.88	474,315
5. Team skills	20.4	7.6	28.0	1.31	0.68	470,760
6. Creative capacities	11.3	8.4	19.7	1.23	0.68	469,467
7. Broad liberal education	14.1	9.5	23.6	1.28	0.71	470,795
8. Communication skills	23.4	18.2	41.6	1.55	0.84	473,065
9. Find, use resources	25.6	10.2	35.8	1.41	0.74	470,526
10. Values development	14.5	6.3	20.8	1.21	0.64	468,492
11. Critical analysis	25.1	20.0	45.1	1.61	0.85	473,929
12. Interest in learning	26.0	9.4	35.4	1.39	0.73	469,334

The percentage of instructors selecting each objective as relevant (i.e., either *Important* or *Essential*) varied substantially, which shows faculty discriminated in their selections. Instructors rated Objectives 1, 2, and 3, respectively, as relevant in over 70% of all classes. In contrast, Objectives 6 and 10 were selected least frequently, identified in one of five classes. These findings are similar to what Hoyt and Lee reported (2002a). The mean number of objectives identified as relevant per class was 5.1 ($SD = 2.6$).

The percentage of total classes in which each objective was selected as relevant is reported by year in Figure 1. There was a slight downward trend in proportion across all objectives, with a small reversal the last two years (2010 and 2011). This is consistent with a declining pattern in the mean number of objectives selected each year, beginning with 5.4 in 2002, dropping to 4.9 in 2010, and increasing slightly to 5.0 in 2011.

Figure 1

Percentage of Classes Identifying Each Learning Objective as Relevant by Year



Year	Obj 1	Obj 2	Obj 3	Obj 4	Obj 5	Obj 6	Obj 7	Obj 8	Obj 9	Obj 10	Obj 11	Obj 12
2002	79.6%	76.8%	77.8%	57.0%	33.4%	23.4%	27.1%	47.0%	42.0%	24.5%	49.2%	42.1%
2003	80.4%	76.6%	78.0%	57.9%	33.8%	23.2%	26.3%	47.2%	42.3%	25.4%	50.4%	42.9%
2004	81.5%	78.1%	78.8%	58.7%	34.5%	23.4%	27.2%	47.3%	43.1%	26.4%	51.6%	43.8%
2005	80.1%	76.4%	76.7%	55.6%	32.3%	21.1%	25.8%	45.7%	40.6%	24.8%	49.2%	41.1%
2006	79.7%	75.3%	75.9%	54.8%	30.5%	21.1%	24.5%	44.2%	38.6%	23.2%	47.3%	38.4%
2007	79.1%	74.8%	76.4%	52.6%	29.3%	20.9%	24.8%	44.4%	37.3%	22.2%	47.5%	37.6%
2008	78.3%	74.1%	75.4%	52.4%	28.2%	20.4%	24.2%	42.1%	36.3%	20.6%	46.1%	35.6%
2009	77.1%	72.8%	74.1%	50.7%	27.5%	20.2%	24.3%	41.3%	35.6%	20.0%	45.3%	35.3%
2010	76.4%	72.8%	74.6%	51.4%	27.9%	20.0%	24.4%	42.0%	35.5%	20.5%	45.9%	35.5%
2011	76.9%	73.0%	74.5%	51.8%	28.2%	20.3%	24.6%	43.1%	36.5%	21.6%	45.9%	35.7%

Summary

Among the 13 learning objectives in the *Diagnostic Feedback* instrument, six retained the wording used in previous editions, two were reworded to reflect changes in the field, and one incorporated wording from two that were conceptually related. Four new ones were added to address AAC&U VALUE rubrics. Of the 19 teaching methods in the revised instrument, six first appeared in 1969 and one was added to the forms produced in 1972, 1975, and 1988. The 1998 version included nine new teaching methods. The new instrument's six student and course characteristics included three carried over from previous editions, one adopted from the *Short Form*, one that combined two that were conceptually related, and one new objective. One overall summary rating pertaining to attitude change was dropped, whereas global ratings of the teacher and course were retained. Instructors were discriminating in their selection of objectives. They rated Objectives 1, 2, and 3 as relevant in over 70% of all classes, whereas they selected Objectives 6 and 10 only about 20% of the time.

Variables in the Adjusted Score Models

Mean scores on several items pertaining to student and course characteristics are used in computing adjusted scores on student ratings of learning objectives and the two overall summary measures. The adjustments control for extraneous factors that can affect ratings (e.g., student work habits, motivation, effort). They are intended to "level the playing field" between instructors who teach highly motivated students with good work habits and those whose students may be lacking in those characteristics.

On the DF, students answered questions about three course characteristics:

33. *Amount of reading*
34. *Amount of work in other (non-reading) assignments*
35. *Difficulty of subject matter.*

They used the scale, 1 = *Much Less than Most Courses*, 2 = *Less than Most Courses*, 3 = *About Average*, 4 = *More than Most Courses*, and 5 = *Much More than Most Courses*. They also answered questions about the following attitudes and behaviors in the course:

36. *I had a strong desire to take this course.*
37. *I worked harder on this course than on most courses I have taken.*
38. *I really wanted to take a course from this instructor.*
39. *I really wanted to take this course regardless of who taught it.*
40. *As a result of taking this course, I have more positive feelings toward the field of study.*
43. *As a rule, I put forth more effort than other students on academic work.*

Students responded to the questions above using the same scale reported previously for the two overall summary measures. On the SF, students were asked to respond to only Items 39 ("Course Motivation") and 43 ("Work Habits").

Responses to Items 39 and 43 are the two extraneous variables that historically carry the greatest weight in the regression models for computing adjusted scores (Hoyt & Lee, 2002a). Items 33 and 34 are used in the models for creating a residual score on course difficulty (Item 35), which probably reflects differences among disciplines (Hoyt & Lee, 2002a) after controlling for the instructor's influence. Items 33 and 34 also are used in creating a residual student effort (Item 37) score, which probably reflects student background, after removing their influence. Items 36 and 38 were included to create a residual score on "Other Motivation," but that residual added so little variance to the regression models that Hoyt and Lee (2002a) did not include it in any adjusted score formulas. Therefore, Items 36 and 38 will not be retained in the updated instrument.

As mentioned previously, Hoyt and Lee (2002a) created two residual scores for the adjusted score formulas they defined as D_N and E_N. The two measures represent the average student perception of course difficulty and effort after the instructor's influence is removed. The instructor's influence is measured in mean responses to Items 33 (amount of reading), 34

(amount of work in non-reading assignments), and 8 (stimulating intellectual effort). The formulas for predicting difficulty were:

Predicted X_{35} = .087(X_8) + .222(X_{33}) + .490(X_{34}) + .675; R^2 = .427
D_N = Mean of X_{35} – Predicted X_{35}

For effort, the formulas were:

Predicted X_{37} = .371(X_8) + .089(X_{33}) + .535(X_{34}) + -.006; R^2 = .618
E_N = Mean of X_{37} – Predicted X_{37}

Both formulas show that the more reading and "other work" is required, and the more students perceive the instructor stimulates their intellectual effort, the more difficult they perceive the course to be and the more effort they report putting forth. Table 25 presents regression coefficients and constants for replicating the modeling Hoyt and Lee (2002a) performed on each objective and overall measures. The coefficients and R^2 values are similar in magnitude to those reported previously (see Hoyt & Lee, 2002a, p. 37). Namely, student work habits (WH) and motivation (CM) are generally the most potent predictors. Classes comprised primarily of students who report good work habits and a strong desire to take the course regardless of who taught it tend to receive higher ratings.

Table 25

Regression Coefficients and Constants for Adjusting Ratings On the Diagnostic Form

		Regression Coefficient						
Criterion	Constant	CM	WH	N	D_N	E_N	$1 + R^2$	Grand M
21. Factual knowledge	1.635	.301	.387	.00	.114	-.145	1.235	4.1528
22. Principles/theories	1.625	.295	.381	--	.100	-.161	1.220	4.0922
23. Applications	1.568	.297	.415	-.002	-.092	-.138	1.253	4.1320
24. Prof skill, viewpoints	1.244	.333	.459	-.002	-.111	-.095	1.298	4.0942
25. Team skills	0.226	.314	.642	-.003	-.419	-.205	1.248	3.6807
26. Creative Capacities	0.500	.269	.611	-.007	-.607	-.119	1.311	3.5980
27. Broad liberal education	0.514	.313	.550	-.003	-.307	-.186	1.215	3.6147
28. Communication skills	0.665	.200	.634	-.006	-.588	-.175	1.290	3.6328
29. Find, use resources	0.819	.224	.597	-.003	-.201	-.196	1.241	3.7974
30. Values development	0.534	.301	.565	-.001	-.424	-.130	1.269	3.6998
31. Critical analysis	1.181	.210	.522	-.002	-.237	-.164	1.189	3.8523
32. Interest in learning	1.030	.310	.493	-.003	-.126	-.141	1.256	3.9124
41. Excellent teacher	2.344	.294	.253	-.001	-.094	-.219	1.117	4.2989
42. Excellent course	1.277	.512	.276	-.001	-.167	.010	1.324	4.0670

Note. CM = Course Motivation (Item 39), WH = Work Habits (Item 43), N = enrollment, D_N = Difficulty unrelated to the instructor, E_N = Effort unrelated to the instructor. Classes with response rates less than 75% or not reporting the number enrolled were excluded.

Relative to "Course Motivation" and "Work Habits," neither residual score (i.e., D_N and E_N) adds much to the adjusted score formulas, which is consistent with previous findings (Hoyt & Lee, 2002a). Given the small contributions the residual difficulty and effort scores made to the regression models, the following investigation was undertaken to identify another potential variable for the adjusted score formulas.

Student Background

In Spring 2011, IDEA research staff examined the utility of adding the IDEA SF item *My background prepared me well for this course's requirements* (14SF) to the formulas used in adjusting ratings. Decades of research in cognitive psychology and education have demonstrated convincingly the critical role of background knowledge in student learning. Consequently, professional staff set out to test its influence on student ratings.

Hoyt and Lee (2003) had performed a preliminary analysis to determine whether 14SF should be retained and added to the DF. They compared the amount of additional variance explained in regression models when 14SF was included as an extraneous variable. Adding 14SF to the models explained significant additional variance in mean student progress ratings on each of the 12 learning objectives (when including only classes where the instructor rated the objective as relevant) and the two overall summary measures.

For the 2011 analysis, the research questions were:

1. Do the IDEA SF items currently used as extraneous variables in adjusted scores (i.e., 13SF, *As a rule, I put forth more effort than other students on academic work*, and 15SF, *I really wanted to take this course regardless of who taught it*) explain significant variance in 14SF? If the amount of variance explained in 14SF is large, then not much would be gained by adding 14SF as a covariate.

2. Do the intercepts and slopes produced by regression models that include 13SF, 14SF, and 15SF differ significantly by discipline groups? That is, do the effects of the three explanatory variables on the models depend upon disciplinary groupings?

3. If the answer to Question 2 is yes, then can clusters of discipline groups be formed or do all disciplines fit into one cluster?

4. If the disciplines fall into distinct clusters, what are the magnitudes of the correlations between outcome measures adjusted on the basis of average coefficients in separate clusters versus those adjusted on average coefficients in the overall database? What is the magnitude of the difference in means and variances between adjusted scores produced under those two conditions?

Procedures. We analyzed SF student ratings collected from 2006 to 2010 ($N = 134{,}068$) because we wanted to include data from the previous five years at the time the study was conducted. Historical exclusions described previously in this report were employed with the exception that only classes using the SF were included. To answer the first research question, we

regressed 14SF (student background) on 13SF (work habits) and 15SF (motivation). Tables 26 and 27 report the results of that analysis. The full model produced an $R^2 = .33$, which left 67% of the variance in 14SF unexplained. Because 13SF and 15SF explained only about one-third of the variance in 14SF, it was possible much could be gained by adding 14SF to the adjusted score regression models. We, therefore, examined whether the intercepts and slopes produced by regression models that included 13SF, 14SF, and 15SF differed significantly by discipline groups.

Table 26

Results of Student Background Regressed on Work Habits and Motivation

Source	*df*	*SS*	*MS*	*F*	*p*
Model	2	8649.13	4324.57	33437.0	<.0001
Error	134065	17339	0.13		
Total	134067	25988			

Table 27

*Summary of Multiple Regression Analysis for Student Background (*N = *134,068)*

Variable	β	SE(β)	*t*	*p*
Motivation	.28	0.002	141.27	<.0001
Work habits	.50	0.003	149.08	<.0001

Note. $R^2 = .33$.

We then formed discipline groupings based on IDEA procedures for determining discipline norms as described in Hoyt and Lee (2002b). We computed the frequencies of the total number of classes for all department codes in the overall IDEA SRI database (SF only). From this information, we rolled some sub-disciplines up into major disciplines or combined similar disciplines to create major groupings that were each comprised of at least 400 classes. This resulted in 39 unique discipline groups.

Following these steps, we then employed the SAS GLM select procedure to conduct multiple regression analyses, separately, on the following student outcome variables: student ratings of progress on each of the IDEA learning objectives (only including classes when the instructor rated the objective as relevant), overall excellence of the instructor (17SF), overall excellence of the course (18SF), and student ratings of progress on objectives the instructor identified as relevant (PRO). Explanatory variables were student work habits (13SF), student background (14SF), student motivation to take the course (15SF), as well as the interactions of each of these with the discipline groupings. In each case, 14SF added significant variance to the regression model, and discipline group interacted significantly with either 14SF or 15SF.

We then produced an output file that contained, for each model, the intercepts and slopes for the interaction terms. Next, we conducted a cluster analysis on those data, employing the agglomerative hierarchical method, which began with as many clusters as the number of disciplines in the dataset ($n = 39$). The most similar disciplines were then grouped first, and these initial groups were merged according to their similarities. Gradually, the similarity among groups decreases, and all disciplines were combined into a single cluster.

Prior to conducting the analysis, we standardized the input variables so that all slopes and intercepts would be equally important in the clustering. Applying Ward's linkage method, the biggest jump in pseudo *t* statistics, an indicator of the number of clusters, was found when advancing from two clusters to one. However, we did not consider this a substantial distinction between clusters. We also employed the PRINCOMP procedure to conduct a principal components analysis on the same data that went into the clustering procedure. In addition, we produced two- and three-dimensional scatterplots of pairs of factors to see how far apart and how distinct the clusters were from one another. We discerned no clear distinction between clusters, and we concluded there was only a single cluster for disciplines. Nevertheless, we performed some additional analyses described in the next paragraph.

We next computed the adjusted scores for each of the 15 outcome measures for the entire data set of 39 disciplines, using the standardized slopes and intercepts averaged across clusters. In addition, we computed adjusted scores averaged within the two clusters. Then, we created plots of adjusted scores using the one cluster with those using two clusters. Using different color schemes for the data based on one versus two clusters, it was found that in most cases there was substantial overlap in the plots. In only three instances were the predictions different, depending on the clusters: student ratings of progress on Objectives 6, 7, and 8. The plots for the remaining outcomes were not different enough to distinguish the clusters. Therefore, we concluded a single cluster was appropriate for all objectives. There was no logical reason as to why the two clusters were formed as they were, and they were not uniquely identified. Most importantly, no clear distinction between one and two clusters was found for student ratings of excellence of the course, excellence of the instructor, and progress on relevant objectives.

Based on the results of the analyses performed on 14SF ("Student Background"), the decision was made to include it in the updated instruments.

Summary

The patterns of regression coefficients for the adjusted score models are similar to those reported previously. Student work habits and motivation remain the most potent predictors. Relative to those variables neither residual score adds much to the adjusted score formulas, which is also consistent with previous findings. Item 14SF ("Student Background") should be included in the updated instruments.

Reliability

Reliability evidence is important for determining whether aggregated student ratings are consistent enough to be used for making administrative decisions about teaching effectiveness. If average ratings changed dramatically from one class to the next for a given instructor, then summative decisions would be suspect. A second question concerns the amount of standard error typically found in mean ratings of specific items. Such information can help inform administrators about an expected range of values when looking at a specific mean score.

Hoyt and Lee (2002a) computed split-half reliabilities on each of the 47 items on the Diagnostic Form. They then applied the Spearman-Brown formula to estimate reliabilities for class size ranges of 10-14, 15-34, 35-49, and 50+. They were able to demonstrate adequate reliability of student ratings at the class level, which is a pre-requisite of instructor-level reliability (Gillmore, 2000). However, it is possible to have adequate reliability at the class level (consistency in ratings by students in the same class) without having it at the instructor level (consistency in ratings of the same instructor across different classes). We, therefore, computed inter-class reliability coefficients to obtain a measure of instructor-level reliability.

The following steps were taken to create a database for reliability analysis.

1) A unique identifier was created for each instructor and course.

    a. The unique identifier for the instructor was a concatenation of the following data elements: identity fkey (unique identifier for the institution), dept_code, last_name, first_name, and middle_name. Although this was imperfect as a unique identifier, it was as close as possible with the available data elements. (For example, it could be possible to have two instructors with the name J. Smith in the same department in the same institution.)

    b. The unique identifier for the course was a concatenation of the following data elements: dept_code and course_num.

2) The number of courses in the remaining database that belong to a specific instructor was determined, and only instructors with five or more classes in the database were retained.

3) The remaining instructors were each assigned a unique random number that was then sorted into ascending order, and the first 2500 were selected for the reliability analysis.

4) Classes taught by those 2500 instructors were matched against the datasets containing individual student responses to create a dataset that was input into the reliability analysis. This dataset contained all student responses for all classes taught by the 2500 selected instructors.

The HPMIXED procedure in SAS was applied to produce inter-class correlation coefficients for all 47 items of the DF for the classes taught by the 2500 instructors. We then applied the Spearman-Brown prophecy formula to estimate reliabilities for 1 to 15 classes. We

plotted inter-class reliability coefficients by number of classes. The magnitudes of the coefficients increased as the number of classes rated increased. See Appendix D for plots of coefficients for the two overall summary measures, which largely reflect those of all items on the DF. All reliabilities approached or exceeded .60 for a single class. When at least two classes were rated, reliability coefficients were above .70 and most approached or exceeded .80. All reliability coefficients were .90 or greater when at least seven classes were rated. Standard errors were approximately .30 or lower for all but one teaching method: Item 5 (*Formed "teams" or "discussion groups" to facilitate learning*). When at least four classes were rated, the SEM for Item 5 was approximately .30. Coefficients and standard errors of measurement (SEM) for all items by number of classes rated appear in Appendix E.

We then computed inter-class reliability coefficients and SEM for all 47 items by class size groupings of 10-14 (small), 15-34 (medium), 35-49 (large), and 50+ (very large). As class size rose, reliability coefficients increased and standard errors decreased. Appendix F presents coefficients and SEMs by class size groupings when one class was rated. We again applied the Spearman-Brown prophecy formula to estimate reliabilities and SEMs when 1 to 15 classes were rated. For small classes all coefficients were .90 or greater when at least nine classes were rated, for medium classes when at least six classes were rated, for large classes when five or more were rated, and for very large classes when at least four were rated.

Teaching Style Scales

Hoyt and Lee (2002a) reported evidence for the reliability of five *a priori* teaching style scales. For the current analysis, we computed internal consistency reliabilities using Cronbach's Alpha. The coefficients shown in Table 28 indicate each scale has high reliability.

Table 28

Internal Consistency Reliabilities for Teaching Style Scales

Scale	Coefficient Alpha
Stimulating student interest	.946
Fostering student collaboration	.854
Establishing rapport	.935
Encouraging student involvement	.876
Structuring classroom experiences	.935

Note: Classes with response rates lower than 75% or not reporting the number enrolled were excluded. Stimulating student interest = DF Items 13, 15, 8, 4; Fostering student collaboration = DF Items 18, 5, 16; Establishing rapport = DF Items 2, 7, 20, 1; Encouraging student involvement = DF Items 9, 14, 19, 11; Structuring classroom experience = DF Items 10, 6, 12, 3, 17.

Revised IDEA SRI Instruments

Based on the analyses and review of information sources described above, IDEA staff decided to update the existing student ratings instruments. The 47-item *Diagnostic Form* will be updated and renamed *Diagnostic Feedback,* and the *Short Form* will be renamed *Learning Essentials*. In addition, a new form that focuses on essential teaching methods will be created called *Teaching Essentials*. The IDEA Student Ratings of Instruction System, therefore, is comprised of the following:

Diagnostic Feedback
- Faculty Information Form
- Diagnostic Feedback
- Diagnostic Report

Learning Essentials
- Faculty Information Form
- Learning Essentials
- Learning Essentials Report

Teaching Essentials
- Teaching Essentials
- Teaching Essentials Report

The item-by-item comparisons of the existing IDEA learning objectives and updated objectives can be found in Appendix I. Appendix J displays the same kind of comparisons for teaching methods.

Diagnostic Feedback

Diagnostic Feedback provides comprehensive student feedback about student progress on relevant course objectives and overall ratings of the instructor and the course. As such, it continues the long-standing IDEA tradition of a two-form system. Instructors complete the *Faculty Information Form,* and students respond to questions on the *Diagnostic Feedback* instrument. The *Diagnostic Report* provides both formative and summative feedback, controls for extraneous factors (e.g., student work habits and motivation) beyond the instructor's control, and reports comparative scores.

Faculty Information Form (FIF). All instructors are required to complete an FIF for every section taught. They begin by rating the relevance of 13 learning objectives, using the scale *Essential, Important*, and *Minor or no Importance*. Then they answer contextual questions about the course, including how frequently they employed each of six different teaching approaches (e.g., lecture, discussion, group-based learning). Next, they indicate the extent to which each of eight academic skills was required of students (e.g., writing, oral communication, mathematical/quantitative work). They are also asked to rate whether each of nine circumstances had a positive, negative, or neutral impact on student learning (e.g., technical support, desire to teach the course, student level of effort). They also identify the principal type of student enrolled in the course (e.g., first-year/sophomores seeking to fulfill a general education or distribution

requirement, graduate or professional students). Finally, they describe the instructional delivery (i.e., face-to-face, online, hybrid).

Diagnostic Feedback. The *Diagnostic Feedback* instrument is comprised of 40 items. First, students indicate how frequently they observed each of 19 teaching methods practiced by the instructor. They then rate the amount of progress they made on the same 13 learning objectives their instructor rated for relevance to the course. They also describe the relative amount of course work and difficulty of the subject matter. Four questions address student characteristics: motivation to take the course, self-efficacy, typical work habits, and adequacy of background. Finally, students rate the overall quality of the teacher and the course.

Learning Essentials

Learning Essentials provides summative feedback about average student progress on relevant learning objectives and overall ratings of the instructor and the course. No formative feedback is provided. Instructors complete the same FIF as is done for *Diagnostic Feedback*. Students respond to the *Learning Essentials* instrument. The 18-item survey includes 13 learning objectives, three student characteristics, and two overall summary measures (excellence of instructor and excellence of course). The *Learning Essentials Report* provides summative feedback, controls for extraneous factors beyond the instructor's control, and reports comparative scores.

Teaching Essentials

As a new instrument, *Teaching Essentials* (TE) is useful for providing formative feedback about essential teaching methods highly correlated with student ratings of the instructor and the course. It can also provide summative feedback focused on student impressions of the instructor and the course. However, no information is collected about student ratings of progress on course objectives, which IDEA believes is the best single measure of teaching effectiveness. Consequently, the TE is a single-form system—instructors do not have to complete an FIF. The *Teaching Essentials* instrument is comprised of 12 items, including seven teaching methods, three student characteristics, and two overall summary measures. As with the other two forms, the *Teaching Essentials Report* controls for extraneous factors beyond the instructor's control and reports comparative scores on excellence of the instructor and the course.

References

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1999). Standards for educational and psychological testing. Washington, DC: American Educational Research Association.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (2014). Standards for educational and psychological testing. Washington, DC: American Educational Research Association.

American Federation of Teachers. (2010). Survey of part-time and adjunct higher education faculty. *American Academic* (Vol. 2). AFT Higher Education.

Association of American Colleges and Universities. (2014). VALUE: Valid Assessment of Learning in Undergraduate Education. http://www.aacu.org/value/rubrics/index_p.cfm

Benton, S. L., & Cashin, W. E. (2011). *IDEA Paper No. 50: Student ratings of teaching: A summary of research and literature*. Manhattan, KS: The IDEA Center. http://theideacenter.org/research-and-papers/idea-papers/50-student-ratings-teaching-summary-research-and-literature

Benton, S. L., & Cashin, W. E. (2014). Student ratings of instruction in college and university courses. In Michael B. Paulsen (Ed.), *Higher Education: Handbook of Theory & Research*, Vol. 29 (pp. 279-326). Dordrecht, The Netherlands: Springer.

Benton, S. L., Duchon, D., & Pallett, W. H. (2013). Validity of self-reported student ratings of instruction. *Assessment & Evaluation in Higher Education, 38,* 377-389.

Braskamp, L. A., & Ory, J. C. (1994). *Assessing faculty work: Enhancing individual and institutional performance.* San Francisco, CA: Jossey-Bass.

Feldman, K. A. (1989). The association between student ratings of specific instructional dimensions and student achievement: Refining and extending the synthesis of data from multisection validity studies. *Research in Higher Education, 30,* 583-645.

Finkel, D. L. (2000). *Teaching with your mouth shut.* Portsmouth, NH: Boynton/Cook Publishers.

Gillmore, G. M. (2000). *Drawing inferences about instructors: the inter-class reliability of student ratings*. University of Washington, Seattle, WA: Office of Educational Assessment.

Hativa, N. (2013a). *Student ratings of instruction: A practical approach to designing, operating, and reporting*. Oron Publications.

Hativa, N. (2013b). *Student ratings of instruction: Recognizing effective teaching*. Oron Publications.

Hativa, N., Barak, R., & Simhi, E. (2001). Exemplary university teachers: Knowledge and beliefs regarding effective teaching dimensions and strategies. *Journal of Higher Education, 72*, 699-729.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). Bayesian model averaging: A tutorial with discussion. *Statistical Science*, 14:382-417.

Hoyt, D. P. (1973a). Identifying effective instructional procedures. *Improving College and University Teaching, 21*, 73-76.

Hoyt, D. P. (1973b). Measurement of instructional effectiveness. *Research in Higher Education, 1,* 367-378.

Hoyt, D. P., & Cashin, W. E. (1977). *IDEA Technical Report No. 1: Development of the IDEA system*. Kansas State University, Manhattan, KS: The IDEA Center.

Hoyt, D. P., Chen, Y., Pallett, W. H., & Gross, A. B. (1999). *IDEA Technical Report No. 11: Revising the IDEA system for obtaining student ratings of instructors and courses*. Kansas State University, Manhattan, KS: The IDEA Center.

Hoyt, D. P., & Lee, E. (2002a). *IDEA Technical Report No. 12: Basic data for the revised IDEA system.* Kansas State University, Manhattan, KS: The IDEA Center.

Hoyt, D. P., & Lee, E. (2002b). *IDEA Technical Report No. 13: Disciplinary differences in student ratings.* Kansas State University, Manhattan, KS: The IDEA Center.

Hoyt, D. P., & Lee, E. (2003, April). *Short form extraneous variable analysis*. Unpublished manuscript. The IDEA Center.

Linn, R. L., & Gronlund, E. (2000). *Measurement and assessment in teaching* (8th ed.). Upper Saddle River, NJ: Merrill.

Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, *11*, 253-388.

Murray, H. G. (1997). Effective teaching behaviors in the college classroom. In R. P. Perry & J. C. Smart (Eds.), *Effective teaching in higher education: Research and practice* (pp. 171-204). New York, NY: Agathon Press.

Schwartz, B. M., & Gurung, R. A. R. (2012). *Evidence-based teaching for higher education*. Washington, DC: American Psychological Association.

Svinicki, M. & McKeachie, W. J. (2011). *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers*. (13th ed.). Belmont, CA: Wadsworth.

Appendix A: Pearson *r* Correlations Among Student Ratings of 47 Items on Diagnostic Form

Item	SR1	SR2	SR3	SR4	SR5	SR6	SR7	SR8	SR9	SR10	SR11	SR12	SR13	SR14	SR15	SR16	SR17	SR18	SR19	SR20	SR21	SR22
SR1	—																					
SR2	.90	—																				
SR3	.75	.79	—																			
SR4	.84	.85	.77	—																		
SR5	.46	.49	.43	.40	—																	
SR6	.82	.85	.77	.91	.44	—																
SR7	.80	.83	.74	.77	.51	.79	—															
SR8	.77	.83	.74	.80	.45	.79	.80	—														
SR9	.61	.65	.60	.62	.55	.62	.67	.69	—													
SR10	.81	.85	.78	.85	.35	.89	.77	.75	.59	—												
SR11	.69	.71	.60	.80	.42	.79	.64	.66	.57	.71	—											
SR12	.67	.71	.76	.74	.28	.76	.64	.67	.49	.76	.62	—										
SR13	.82	.85	.73	.88	.46	.88	.79	.83	.67	.84	.82	.71	—									
SR14	.57	.59	.54	.56	.71	.58	.63	.57	.71	.49	.60	.42	.63	—								
SR15	.80	.84	.74	.79	.55	.79	.84	.87	.74	.75	.68	.63	.83	.73	—							
SR16	.66	.69	.56	.64	.66	.66	.70	.65	.71	.60	.69	.45	.76	.70	.74	—						
SR17	.67	.69	.72	.67	.30	.69	.69	.65	.48	.72	.56	.71	.65	.40	.63	.48	—					
SR18	.74	.78	.67	.67	.73	.70	.77	.73	.66	.65	.61	.55	.73	.70	.81	.78	.60	—				
SR19	.66	.70	.67	.65	.61	.66	.74	.71	.75	.62	.59	.56	.74	.75	.79	.78	.53	.74	—			
SR20	.78	.76	.68	.69	.44	.70	.74	.74	.62	.67	.60	.61	.71	.53	.76	.59	.63	.70	.63	—		
SR21	.67	.72	.68	.77	.28	.78	.67	.75	.51	.74	.64	.72	.75	.46	.69	.47	.62	.58	.52	.63	—	
SR22	.68	.74	.68	.76	.31	.77	.68	.77	.52	.74	.66	.71	.76	.48	.71	.51	.63	.61	.55	.64	.91	—
SR23	.75	.80	.73	.81	.43	.80	.75	.79	.61	.76	.72	.69	.80	.61	.81	.63	.64	.71	.67	.69	.83	.86
SR24	.73	.78	.70	.79	.43	.79	.76	.78	.62	.74	.69	.66	.78	.63	.81	.62	.62	.70	.67	.68	.83	.84
SR25	.52	.57	.49	.49	.84	.51	.58	.55	.59	.45	.49	.37	.54	.75	.66	.65	.39	.76	.63	.51	.45	.49
SR26	.55	.60	.53	.53	.56	.54	.69	.60	.68	.54	.45	.38	.63	.66	.72	.73	.42	.67	.82	.52	.45	.48
SR27	.58	.63	.53	.58	.46	.61	.67	.65	.60	.60	.47	.45	.70	.53	.69	.69	.47	.64	.69	.53	.56	.57
SR28	.56	.61	.53	.55	.60	.56	.67	.63	.72	.54	.51	.40	.65	.63	.69	.79	.44	.67	.79	.54	.46	.49
SR29	.61	.69	.63	.63	.50	.63	.69	.73	.81	.62	.55	.55	.67	.65	.76	.65	.54	.69	.71	.65	.65	.67

SR30	.63	.69	.58	.67	.52	.67	.69	.69	.67	.64	.65	.51	.75	.62	.76	.80	.51	.71	.71	.59	.60	.65
SR31	.63	.70	.60	.66	.51	.67	.70	.75	.70	.64	.62	.53	.75	.58	.74	.77	.52	.69	.75	.62	.62	.68
SR32	.72	.80	.68	.74	.48	.74	.75	.81	.68	.72	.65	.62	.80	.61	.82	.72	.61	.75	.72	.69	.75	.78
SR33	.02	.06	.06	.10	.12	.10	.05	.24	.22	.03	.14	.08	.15	.03	.09	.21	.04	.08	.15	.11	.14	.14
SR34	.10	.13	.23	.08	.18	.05	.19	.31	.25	-.01	-.04	.12	.04	.24	.28	.04	.08	.19	.26	.20	.17	.17
SR35	-.05	.00	.01	.00	-.12	-.03	.00	.26	-.02	-.10	-.09	.08	-.03	-.12	.06	-.20	.02	-.03	-.06	.10	.20	.19
SR36	.44	.46	.37	.50	.24	.49	.44	.47	.34	.43	.46	.35	.54	.40	.50	.40	.32	.42	.40	.38	.55	.53
SR37	.26	.32	.32	.31	.16	.28	.34	.54	.31	.21	.18	.28	.30	.25	.45	.17	.25	.30	.31	.35	.45	.44
SR38	.69	.72	.59	.68	.36	.69	.69	.70	.53	.67	.60	.55	.72	.51	.72	.55	.53	.63	.56	.65	.67	.67
SR39	.27	.29	.24	.33	.19	.32	.30	.31	.24	.27	.30	.23	.36	.32	.35	.27	.21	.30	.28	.24	.41	.40
SR40	.70	.72	.63	.77	.36	.77	.67	.69	.54	.72	.69	.61	.80	.55	.72	.60	.56	.61	.60	.59	.76	.75
SR41	.85	.87	.77	.84	.36	.85	.77	.77	.57	.90	.69	.74	.84	.49	.76	.59	.71	.67	.63	.71	.74	.74
SR42	.74	.78	.70	.80	.36	.81	.72	.74	.55	.80	.69	.68	.83	.53	.75	.60	.63	.64	.63	.63	.78	.77
SR43	.24	.28	.25	.27	.25	.28	.34	.37	.30	.19	.26	.20	.30	.34	.39	.28	.20	.34	.31	.33	.38	.37
SR44	.64	.66	.69	.60	.59	.61	.66	.62	.64	.56	.53	.56	.62	.71	.70	.61	.51	.68	.73	.60	.53	.55
SR45	.57	.60	.58	.61	.34	.59	.57	.66	.50	.52	.48	.53	.59	.44	.63	.44	.48	.54	.53	.55	.58	.58
SR46	.58	.61	.59	.63	.34	.60	.61	.74	.53	.53	.48	.52	.60	.46	.70	.45	.49	.55	.56	.59	.62	.61
SR47	.41	.43	.46	.42	.37	.43	.42	.44	.56	.39	.43	.41	.45	.50	.46	.42	.35	.44	.47	.48	.43	.41

Item	SR23	SR24	SR25	SR26	SR27	SR28	SR29	SR30	SR31	SR32	SR33	SR34	SR35	SR36	SR37	SR38	SR39	SR40	SR41	SR42	SR43	SR44	SR45	SR46	SR47
SR23	—																								
SR24	.91	—																							
SR25	.60	.61	—																						
SR26	.61	.64	.66	—																					
SR27	.60	.63	.58	.84	—																				
SR28	.61	.61	.66	.87	.78	—																			
SR29	.75	.74	.67	.71	.67	.76	—																		
SR30	.73	.72	.68	.76	.77	.80	.78	—																	
SR31	.73	.71	.61	.74	.73	.82	.80	.85	—																
SR32	.83	.82	.66	.72	.75	.74	.84	.85	.86	—															
SR33	.07	.06	.08	.09	.15	.26	.20	.20	.31	.17	—														
SR34	.23	.24	.23	.23	.11	.17	.30	.10	.15	.20	.24	—													
SR35	.07	.08	-.07	-.13	-.06	-.11	.08	-.09	.03	.07	.42	.56	—												
SR36	.55	.59	.36	.39	.43	.34	.40	.45	.40	.53	.08	.13	.09	—											
SR37	.43	.45	.28	.28	.27	.26	.42	.29	.35	.43	.34	.67	.67	.47	—										
SR38	.70	.72	.49	.52	.55	.51	.60	.60	.60	.69	.07	.16	.12	.64	.43	—									
SR39	.41	.45	.31	.30	.32	.25	.31	.34	.28	.39	.07	.16	.10	.79	.41	.36	—								
SR40	.79	.80	.49	.55	.60	.54	.60	.66	.63	.74	.09	.10	.01	.78	.42	.74	.60	—							
SR41	.78	.75	.46	.53	.58	.53	.61	.62	.64	.74	.03	.06	-.03	.47	.28	.75	.29	.77	—						
SR42	.80	.81	.49	.57	.63	.56	.62	.66	.65	.76	.06	.08	-.01	.73	.40	.75	.54	.91	.86	—					
SR43	.38	.41	.34	.33	.32	.32	.39	.35	.33	.38	.16	.32	.26	.37	.47	.41	.34	.36	.21	.32	—				
SR44	.65	.64	.63	.62	.53	.61	.65	.58	.58	.63	.08	.36	-.04	.40	.33	.55	.29	.59	.62	.61	.37	—			
SR45	.62	.61	.40	.42	.43	.44	.52	.47	.51	.59	.22	.35	.26	.48	.54	.55	.34	.61	.61	.62	.42	.62	—		
SR46	.65	.65	.43	.47	.47	.48	.57	.51	.56	.62	.26	.45	.36	.45	.65	.57	.32	.60	.60	.61	.46	.60	.81	—	
SR47	.45	.45	.40	.36	.34	.37	.51	.39	.40	.43	.16	.22	.05	.26	.24	.37	.21	.41	.41	.40	.29	.55	.43	.42	—

Appendix B: Bayesian Model Averaging on 12 Learning Objectives

Objective = 1 Class Size = Small

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.67	-.03			.05	-.04	.09		.13		.05		.06	.06	.03	.03	-.06		.04	-.05	.02
3 vars	.65						.16		.16				.08								
4 vars	.65					-.04	.17		.17				.08								
5 vars	.66						.14		.15				.07	.10			-.07				
6 vars	.66						.11		.15		.06		.06	.08			-.07				
7 vars	.66				.05		.09		.14		.05		.06	.07			-.07				
8 vars	.67						.10		.13		.05		.07	.08		.05	-.06			-.04	

Objective = 1 Class Size = Medium

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.73	-.04			.04	-.04	.11		.14	-.01	.03		.08	.08	.04	.01	-.06		.04	-.06	.03
3 vars	.70						.16		.15				.10								
4 vars	.71						.18		.18				.10							-.06	
5 vars	.71						.15		.15				.08	.10			-.08				
6 vars	.72						.14		.16				.09	.11			-.06			-.05	
7 vars	.72						.13		.14				.09	.10		.04	-.06			-.05	
8 vars	.72					-.03	.14		.16				.09	.09	.04		-.05			-.06	

Objective = 1 Class Size = Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.76	-.06			.03	-.04	.13	.02	.16		.04		.06	.06	.03		-.05		.04	-.06	.03
3 vars	.74						.26		.19								-.07				
4 vars	.75						.18		.20				.09							-.07	
5 vars	.75					-.03	.19		.20				.08							-.05	
6 vars	.75					-.04	.18		.20				.08		.03					-.06	
7 vars	.75					-.02	.17		.19				.07	.07			-.04			-.04	
8 vars	.75					-.03	.16		.19				.07	.06	.03		-.04			-.05	

Objective = 1 Class Size = Very Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.81	-.07		-.02		-.04	.18		.15			.02	.08	.05	.03	.05	-.06	.02	.04	-.07	.03
3 vars	.78						.27		.19								-.08				
4 vars	.79						.19		.20				.08							-.08	
5 vars	.79						.21		.20				.07				-.04			-.06	
6 vars	.79	-.05					.20		.18				.10			.07				-.10	
7 vars	.80	-.05					.22		.16				.09			.09	-.04			-.07	
8 vars	.80	-.05					.18		.15				.08	.07		.09	-.05			-.07	

Objective = 2 Class Size = Small

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.67	-.05	.02		.05	-.03	.07		.13	-.02	.05		.06	.07	.01	.03	-.03		.03	-.03	.03
3 vars	.65						.16		.17				.08								
4 vars	.66						.11		.16		.07		.07								
5 vars	.66						.09		.15		.06		.07	.05							
6 vars	.67						.09		.15		.06		.06	.08			-.04				
7 vars	.67						.09		.13		.06		.06	.08		.04	-.05				
8 vars	.67						.09		.14		.05		.07	.08		.05	-.03			-.03	

Objective = 2 Class Size = Medium

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.72	-.06	.03		.04	-.04	.08		.15	-.03	.02	.02	.07	.07	.03	.02	-.03		.04	-.04	.03
3 vars	.70						.14		.18				.10								
4 vars	.71					-.03	.16		.19				.09								
5 vars	.71						.11		.17				.10	.08						-.05	
6 vars	.71						.11		.18	-.03			.09	.09						-.04	
7 vars	.72						.11		.17	-.03			.10	.08		.04				-.05	
8 vars	.72					-.03	.11		.16				.08	.09			-.04		.05	-.03	

Objective = 2 Class Size = Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.75	-.06	.03			-.02	.08	.02	.16	-.02	.03	.04	.07	.05			-.03		.05	-.03	.03
3 vars	.73						.13		.18				.10								
4 vars	.74						.15		.20				.10							-.04	
5 vars	.74						.11		.19				.10	.06						-.05	
6 vars	.74						.10		.17				.10	.06						-.05	.03
7 vars	.74						.11		.18	-.03			.09	.06						-.04	.03
8 vars	.74	-.05	.06				.10		.16				.09	.06						-.06	.04

Objective = 2 Class Size = Very Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.81	-.07				-.03	.08		.15	-.03	.04	.06	.07	.03		.05	-.04		.07	-.02	.03
3 vars	.79						.14		.18				.09								
4 vars	.79						.15		.19				.09							-.03	
5 vars	.80						.12		.19			.06	.08				-.04				
6 vars	.80						.07		.19		.05	.06	.07				-.04				
7 vars	.80					-.04	.11		.17			.07	.08				-.04		.07		
8 vars	.80	-.04				-.05	.13		.17			.07	.08				-.04		.08		

Objective = 3 Class Size = Small

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.73	-.04	.05	.01	.08	-.03	.05	.01	.06	-.03	.03	.06	.02	.02	.06	.08	-.04		.04	.01	.01
3 vars	.71				.13		.12									.17					
4 vars	.71				.11		.11					.06				.17					
5 vars	.72				.09		.10		.07			.06				.12					
6 vars	.72		.06		.08		.08		.06			.06				.11					
7 vars	.72		.05		.07		.07		.05			.06	.04			.11					
8 vars	.73		.05		.08		.07		.06			.05	.04		.04	.09					

Objective = 3 Class Size = Medium

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.78	-.05	.05	.01	.08	-.04	.05		.07	-.03	.02	.06	.03		.06	.09	-.03		.06	.02	.01
3 vars	.75				.18								.08			.19					
4 vars	.76				.13							.07	.08			.18					
5 vars	.77						.11		.08			.08	.06			.13					
6 vars	.77				.07		.07		.07			.06	.05			.13					
7 vars	.77				.08		.07		.07			.06	.05			.11			.03		
8 vars	.77				.08		.07		.07			.07	.05			.12	-.04		.05		

Objective = 3 Class Size = Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.79	-.07	.09		.08	-.04	.05		.09	-.02		.10	.03	-.06	.03	.13	-.02		.07		
3 vars	.76						.18		.12							.14					
4 vars	.77								.13			.12	.08			.15					
5 vars	.78		.07						.11			.10	.06			.12					
6 vars	.78	-.07	.13		.10				.09			.08				.12					
7 vars	.78	-.07	.11		.08				.09			.08	.05			.13					
8 vars	.78	-.07	.12		.10				.10			.10	.05	-.06		.14					

Objective = 3 Class Size = Very Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.80	-.09	.07		.06	-.06	.08		.10			.13	.03	-.10	.03	.16			.07		
3 vars	.78								.15			.16				.16					
4 vars	.79						.09		.12			.11				.15					
5 vars	.79						.13		.13			.13		-.08		.16					
6 vars	.79					-.02	.13		.12			.13		-.08		.18					
7 vars	.80	-.06				-.05	.11		.11			.12				.16			.07		
8 vars	.80	-.08	.11			-.03	.12		.11			.14		-.09		.18					

Objective = 4 Class Size = Small

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.72	-.03	.02		.08	-.04	.06	.04	.05	-.04	.03	.04	.02	.02	.09	.11	-.05	-.01	.05		.01
3 vars	.69				.12		.12									.19					
4 vars	.70				.12		.12								.05	.16					
5 vars	.70				.12	-.04	.12								.08	.16					
6 vars	.71				.10	-.04	.12		.06						.08	.13					
7 vars	.71				.11	-.04	.12		.07	-.04					.10	.13					
8 vars	.71				.10	-.04	.10	.04	.07	-.04					.10	.11					

Objective = 4 Class Size = Medium

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.77	-.05	.02		.09	-.06	.08	.04	.06	-.04	.02	.04	.03		.10	.10	-.03	-.01	.06		.02
3 vars	.74				.11		.14									.19					
4 vars	.75				.11		.13								.04	.16					
5 vars	.75				.10	-.05	.14								.08	.17					
6 vars	.76				.09	-.05	.13		.07						.09	.11					
7 vars	.76				.09	-.05	.13		.08	-.04					.10	.12					
8 vars	.76				.09	-.07	.12		.08	-.04					.10	.10			.05		

Objective = 4 Class Size = Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.81	-.06			.09	-.07	.11	.03	.09			.05	.02	-.05	.07	.13			.07	-.04	.02
3 vars	.80						.19		.09							.17					
4 vars	.80						.15		.10			.06				.16					
5 vars	.80					-.05	.19		.10						.07	.15					
6 vars	.80					-.08	.18		.10						.07	.12			.06		
7 vars	.81				.07	-.07	.12		.09						.07	.11			.06		
8 vars	.81	-.04			.09	-.08	.13		.08						.06	.12			.07		

Objective = 4 Class Size = Very Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.78	-.10			.07	-.07	.09	.04	.09			.07			.06	.14			.09	-.05	
3 vars	.77						.18		.08							.18					
4 vars	.77						.11		.10			.09				.17					
5 vars	.77	-.06					.14		.10			.09				.19					
6 vars	.78	-.06					.14		.10			.09				.21				-.03	
7 vars	.78	-.07					.13	.05	.10			.10				.19				-.04	
8 vars	.78	-.07				-.07	.14		.09			.08			.04	.16			.08		

Objective = 5 Class Size = Small

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.66	-.07		-.03		.35	-.03		.05	-.03	.07		-.02		.16	.17	-.05		.13	-.06	
3 vars	.64					.36									.13	.18					
4 vars	.64					.37									.15	.22				-.07	
5 vars	.65					.34									.15	.17			.09	-.08	
6 vars	.65	-.07				.34									.15	.20			.12	-.08	
7 vars	.65	-.06				.34									.15	.20	-.05		.13	-.07	
8 vars	.66	-.08				.34					.04				.15	.19	-.05		.13	-.07	

Objective = 5 Class Size = Medium

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.71	-.07		-.04		.36			.05	-.05	.06	-.02			.19	.14	-.05	-.01	.11	-.03	.02
3 vars	.70					.37									.16	.15					
4 vars	.70					.39									.16	.18	-.06				
5 vars	.71					.35									.16	.14	-.09		.10		
6 vars	.71	-.05				.35									.16	.16	-.08		.12		
7 vars	.71	-.05				.35				-.04					.17	.18	-.07		.12		
8 vars	.71	-.07				.35				-.04	.03				.17	.17	-.07		.11		

Objective = 5 Class Size = Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.76	-.04		-.02		.31				-.03		-.03	-.04		.16	.24	-.05		.11		
3 vars	.74					.34									.12	.18					
4 vars	.75					.33						-.08			.15	.23					
5 vars	.75					.34						-.06	-.04		.14	.24					
6 vars	.76					.32							-.07		.15	.20	-.07		.09		
7 vars	.76	-.05				.31							-.05		.15	.22	-.06		.11		
8 vars	.76	-.05				.31				-.03			-.05		.16	.23	-.05		.11		

Objective = 5 Class Size = Very Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.74					.21				-.06		-.06	-.05		.22	.23			.09		
3 vars	.71					.24									.16	.21					
4 vars	.73					.25							-.08		.17	.25					
5 vars	.73					.25				-.06			-.07		.19	.27					
6 vars	.74					.21						-.07	-.06		.20	.21			.09		
7 vars	.74					.21				-.06		-.06	-.05		.22	.23			.09		
8 vars	.74					.21				-.06		-.06	-.05		.22	.23			.09		

Objective = 6 Class Size = Small

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.64							.15	-.03	-.05	.06	-.11	-.08	.09	.05	.20		-.04	.02	.27	
3 vars	.60							.12								.19				.24	
4 vars	.62							.15				-.10				.23				.25	
5 vars	.63							.17				-.08	-.08			.23				.28	
6 vars	.63							.15				-.11	-.09	.10		.20				.27	
7 vars	.63							.15		-.05		-.10	-.08	.10		.21				.28	
8 vars	.64							.15		-.05		-.11	-.08	.10	.05	.19				.27	

Objective = 6 Class Size = Medium

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.71	-.02	-.04			.01		.16	-.05	-.03	.09	-.14	-.10	.08	.02	.21	.04	-.04		.34	
3 vars	.67											-.12				.30				.35	
4 vars	.68							.16					-.13			.17				.35	
5 vars	.69							.17				-.10	-.10			.21				.35	
6 vars	.70							.14			.08	-.11	-.12			.20				.36	
7 vars	.70							.16	-.07			-.13	-.10	.11		.22				.34	
8 vars	.70							.14	-.07		.06	-.13	-.11	.09		.22				.35	

Objective = 6 Class Size = Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.78	-.07	-.10					.19	-.10	-.08	.13	-.26	-.17	.22	.06	.32	.09			.29	
3 vars	.69							.40					-.23							.36	
4 vars	.72							.23					-.21			.22				.30	
5 vars	.74										.21	-.19	-.21			.35				.36	
6 vars	.75							.18				-.23	-.19	.21		.24				.31	
7 vars	.76							.19	-.16			-.24	-.17	.26		.33				.30	
8 vars	.76	-.09						.22	-.17			-.24	-.17	.30		.36				.30	

Objective = 6 Class Size = Very Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.79		-.16					.24			.20	-.29	-.23	.15		.19	.19			.28	-.08
3 vars	.67											-.31				.47	.30				
4 vars	.70							.27				-.33				.27	.27				
5 vars	.71							.31				-.31				.33	.28				-.15
6 vars	.77							.30				-.29	-.25	.29			.15			.27	
7 vars	.78							.26			.14	-.29	-.27	.19			.16			.30	
8 vars	.78							.22			.22	-.27	-.22			.20	.18			.29	-.11

Objective = 7 Class Size = Small

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.54		-.03				.05	.10		-.02	.08	-.14	-.04	.21	-.04	.12	.11			.08	
3 vars	.50							.14						.22		.13					
4 vars	.52							.18				-.14		.31			.14				
5 vars	.53							.13				-.14		.27		.12	.12				
6 vars	.53							.11			.08	-.16		.23		.12	.13				
7 vars	.53							.10			.08	-.15		.22		.10	.10			.05	
8 vars	.54							.10			.08	-.14		.21	-.04	.12	.11			.07	

Objective = 7 Class Size = Medium

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.61	-.03	-.06	-.03	.02	-.01	.09	.10		-.02	.07	-.21		.28	-.03	.09	.15	-.02		.10	
3 vars	.57											-.22		.50			.19				
4 vars	.59							.15				-.21		.40			.16				
5 vars	.60							.13				-.20		.38			.12			.08	
6 vars	.60						.09	.10				-.22		.32			.13			.08	
7 vars	.61						.09	.08				-.22		.31		.05	.13			.07	
8 vars	.61						.09	.08				-.21		.30	-.04	.07	.13			.09	

Objective = 7 Class Size = Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.70		-.17			-.05	.09	.14			.15	-.36	-.08	.36	-.04	.11	.24			.11	-.04
3 vars	.65											-.40		.61			.27				
4 vars	.67							.15				-.39		.50			.24				
5 vars	.67					-.06		.18				-.38		.46			.28				
6 vars	.68		-.18					.18			.15	-.38		.46			.27				
7 vars	.69		-.17					.18			.19	-.37	-.06	.47			.26				
8 vars	.69		-.18					.16			.20	-.35	-.08	.45			.22			.08	

Objective = 7 Class Size = Very Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.77		-.22					.17			.27	-.40	-.11	.45	-.09	.09	.22			.08	
3 vars	.71											-.46		.68			.25				
4 vars	.73							.18				-.46		.57			.20				
5 vars	.74										.25	-.44	-.11	.51			.25				
6 vars	.75		-.21					.23			.22	-.44		.47			.25				
7 vars	.76		-.20					.21			.25	-.42	-.09	.51			.23				
8 vars	.77		-.19					.22			.24	-.41	-.09	.50	-.04		.25				

Objective = 8 Class Size = Small

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.61	-.06		.03		.09		.10	.04	.07	.06	-.04	-.04		-.11	.07	.17		-.04	.20	.04
3 vars	.57							.15									.18			.23	
4 vars	.58							.11	.08								.17			.20	
5 vars	.58							.12	.09			-.07					.20			.20	
6 vars	.59					.07		.14		.10					-.12		.15			.22	
7 vars	.60					.08		.11		.09					-.14	.09	.14			.20	
8 vars	.60					.07		.11		.09		-.04			-.13	.10	.16			.20	

Objective = 8 Class Size = Medium

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.68	-.05	-.02	.02		.07	-.01	.13	.03	.09	.07	-.04	-.05	-.03	-.13	.09	.17		-.05	.29	.01
3 vars	.64							.13									.14			.31	
4 vars	.65							.16				-.09					.19			.31	
5 vars	.66							.12		.11					-.12		.15			.30	
6 vars	.66							.14		.11		-.06			-.09		.18			.30	
7 vars	.67					.05		.15		.12			-.05		-.13		.13			.31	
8 vars	.67					.06		.12		.11			-.06		-.15	.07	.12			.30	

Objective = 8 Class Size = Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.76	-.12				.03		.11	.06	.07	.08	-.09	-.07		-.07	.11	.21		-.06	.33	
3 vars	.71								.09								.16			.34	
4 vars	.73								.14			-.12					.22			.34	
5 vars	.74								.12	.07		-.12					.21			.31	
6 vars	.74	-.10						.13	.12			-.11					.22			.33	
7 vars	.75	-.10						.12	.11	.06		-.10					.21			.30	
8 vars	.75	-.12						.11		.09		-.09			-.09	.14	.20			.33	

Objective = 8 Class Size = Very Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.76	-.14						.19			.08	-.11	-.09	.09		.12	.19		-.12	.36	
3 vars	.71							.20					-.10							.45	
4 vars	.73	-.15						.22									.15			.36	
5 vars	.73	-.14						.25									.19		-.09	.37	
6 vars	.74	-.15						.19								.13	.19		-.12	.35	
7 vars	.75	-.12						.20				-.09				.15	.22		-.13	.34	
8 vars	.75	-.13						.25				-.10	-.09	.16			.18		-.09	.37	

Objective = 9 Class Size = Small

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.70	-.10	.06			-.03	-.02		.09	.25	.06		.03	-.03	.04	.09			.06		.04
3 vars	.68								.10	.27						.15					
4 vars	.68								.10	.26						.11			.06		
5 vars	.69								.08	.26			.03			.11			.05		
6 vars	.69	-.07							.09	.26	.06					.12			.06		
7 vars	.69	-.08							.09	.25	.06					.11			.06		.04
8 vars	.69	-.10	.05						.08	.25	.05					.10			.05		.04

Objective = 9 Class Size = Medium

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.75	-.11	.06		-.02	-.04	-.02		.08	.27	.04	-.02	.04	-.02	.04	.08		.01	.08		.04
3 vars	.73								.15	.29									.09		
4 vars	.73								.10	.28						.09			.06		
5 vars	.73	-.05							.11	.28						.10			.07		
6 vars	.74	-.07							.09	.28			.05			.10			.07		
7 vars	.74	-.10	.07						.08	.28			.04			.09			.06		
8 vars	.74	-.11	.07						.07	.28			.04			.09			.05		.04

Objective = 9 Class Size = Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.82	-.11	.07			-.04			.08	.29	.03		.03	-.10	.03	.14			.09		.03
3 vars	.80								.07	.30						.15					
4 vars	.80								.10	.30				-.06		.18					
5 vars	.81								.09	.30			.04	-.09		.18					
6 vars	.81								.09	.30			.04	-.10		.14			.05		
7 vars	.81	-.09	.10						.07	.30			.04	-.09		.18					
8 vars	.81	-.07				-.04			.08	.31			.05	-.08		.17			.10		

Objective = 9 Class Size = Very Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.78	-.11				-.06			.09	.31					.05	.10		.05	.10		
3 vars	.76								.09	.31						.12					
4 vars	.77	-.09								.32						.23		.07			
5 vars	.77	-.10							.09	.32						.16		.06			
6 vars	.77	-.11							.09	.31						.13		.05	.05		
7 vars	.78	-.11				-.04			.08	.32						.13		.05	.09		
8 vars	.78	-.11				-.06			.09	.31					.05	.10		.05	.10		

Objective = 10 Class Size = Small

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.67	-.07			.04				.08		.04	.06	-.03	.08		.13	.21		.05		
3 vars	.65													.15		.18	.25				
4 vars	.66								.08					.11		.14	.24				
5 vars	.66								.08			.06		.08		.14	.23				
6 vars	.66	-.04							.08			.07		.10		.15	.23				
7 vars	.67	-.05							.08			.07		.09		.14	.21		.04		
8 vars	.67	-.06							.08		.03	.06		.08		.13	.21		.04		

Objective = 10 Class Size = Medium

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.69	-.07	.04		.04		.03	-.02	.05			.07	-.02	.08	-.06	.16	.20	-.02	.03	.02	.02
3 vars	.68													.18		.16	.21				
4 vars	.68								.07					.15		.13	.21				
5 vars	.69								.07			.06		.11		.13	.20				
6 vars	.69								.07			.07		.10	-.05	.16	.21				
7 vars	.69	-.04							.07			.08		.11	-.05	.17	.21				
8 vars	.69	-.04							.06			.08		.11	-.06	.17	.20		.04		

Objective = 10 Class Size = Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.73	-.06			.17	.03							-.04	.10	-.11	.18	.20			.07	
3 vars	.71				.19											.13	.23				
4 vars	.72				.17										-.08	.19	.25				
5 vars	.73				.16										-.10	.17	.23			.07	
6 vars	.73	-.06			.19										-.10	.19	.24			.07	
7 vars	.73	-.06			.15									.07	-.09	.18	.23			.06	
8 vars	.73	-.06			.17								-.04	.09	-.09	.18	.21			.07	

Objective = 10 Class Size = Very Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.72	-.11			.22										-.13	.25	.23			.09	
3 vars	.70				.20											.17	.20				
4 vars	.71				.17										-.09	.24	.23				
5 vars	.71	-.10			.22										-.09	.26	.26				
6 vars	.72	-.11			.22										-.13	.25	.23			.09	
7 vars	.72	-.11			.22										-.13	.25	.23			.09	
8 vars	.72	-.11			.22										-.13	.25	.23			.09	

Objective = 11 Class Size = Small

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.67	-.10	.06			.02		.02	.16	.06	.03		-.02	.11	-.07		.14			.09	.02
3 vars	.65								.24								.19			.10	
4 vars	.66								.19					.09			.16			.09	
5 vars	.66	-.05							.20					.12			.16			.09	
6 vars	.66	-.09	.08						.18					.10			.16			.09	
7 vars	.67	-.05							.18	.06				.13	-.06		.16			.10	
8 vars	.67	-.09	.08						.16	.06				.11	-.06		.16			.09	

Objective = 11 Class Size = Medium

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.73	-.11	.07	-.01		.01		.02	.19	.05	.03		-.01	.08	-.08	-.01	.15			.12	.02
3 vars	.70								.25								.17			.11	
4 vars	.71								.25						-.07		.18			.14	
5 vars	.72								.21					.07	-.07		.16			.14	
6 vars	.72								.20	.05				.08	-.09		.15			.12	
7 vars	.72	-.05							.21	.06				.10	-.08		.16			.12	
8 vars	.73	-.09	.09						.19	.06				.08	-.09		.16			.12	

Objective = 11 Class Size = Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.77	-.12	.11	-.04	.03				.19	.03				.09	-.09		.18			.14	
3 vars	.74								.23					.10			.20				
4 vars	.75	-.08							.25					.14			.21				
5 vars	.76								.21					.08	-.09		.19			.14	
6 vars	.76	-.07							.22					.12	-.09		.19			.14	
7 vars	.77	-.12	.09						.20					.10	-.09		.19			.14	
8 vars	.77	-.11	.11	-.03					.20					.10	-.08		.18			.14	

Objective = 11 Class Size = Very Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.74	-.13	.09		.07				.18		-.05			.13	-.06		.20		-.04	.12	
3 vars	.72								.19					.13			.20				
4 vars	.73	-.10							.22					.18			.22				
5 vars	.73	-.10							.20					.18			.18			.07	
6 vars	.74	-.09							.20					.16	-.06		.20			.11	
7 vars	.74	-.11			.06				.19					.12	-.06		.20			.11	
8 vars	.74	-.13	.05		.06				.18					.11	-.06		.20			.11	

Objective = 12 Class Size = Small

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.71	-.06	.08	.	.	-.03	-.02	.	.12	.03	.04	.	.	.09	.	.10	.06	.	.09	.03	.
3 vars	.69	.	.	.	.	.	.	.	.18	.	.	.	.	.17	.	.	.	.	.14	.	.
4 vars	.70	.	.	.	.	.	.	.	.14	.	.	.	.	.14	.	.12	.	.	.11	.	.
5 vars	.70	.	.	.	.	.	.	.	.14	.	.	.	.	.12	.	.11	.06	.	.08	.	.
6 vars	.70	.	.06	.	.	.	.	.	.13	.	.	.	.	.09	.	.10	.06	.	.07	.	.
7 vars	.71	-.06	.09	.	.	.	.	.	.13	.	.	.	.	.10	.	.11	.06	.	.07	.	.
8 vars	.71	-.06	.09	.	.	-.03	.	.	.13	.	.	.	.	.10	.	.11	.07	.	.09	.	.

Objective = 12 Class Size = Medium

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.78	-.09	.10	-.02	.02	-.04	-.02	-.02	.13	.02	.03	-.02	.02	.09	.	.09	.07	.	.10	.02	.02
3 vars	.75	.	.14	.	.	.	.	.	.24	.	.	.	.	.	.	.	.13	.	.	.	.
4 vars	.76	.	.	.	.	.	.	.	.16	.	.	.	.	.14	.	.10	.	.	.11	.	.
5 vars	.77	.	.	.	.	.	.	.	.17	.	.	.	.	.12	.	.09	.05	.	.08	.	.
6 vars	.77	.	.	.	.	-.04	.	.	.16	.	.	.	.	.11	.	.09	.07	.	.11	.	.
7 vars	.77	-.08	.10	.	.	.	.	.	.15	.	.	.	.	.12	.	.10	.05	.	.07	.	.
8 vars	.77	-.08	.10	.	.	-.05	.	.	.14	.	.	.	.	.10	.	.10	.07	.	.11	.	.

Objective = 12 Class Size = Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.80	-.08	.10	.	.	-.05	.	.	.13	.04	.	-.04	.	.11	-.03	.13	.10	.	.09	.	.
3 vars	.78	.	.	.	.	.	.	.	.24	.	.	.	.	.	.	.13	.14	.	.	.	.
4 vars	.78	.	.09	.	.	.	.	.	.19	.	.	.	.	.	.	.11	.12	.	.	.	.
5 vars	.79	-.06	.14	.	.	.	.	.	.19	.	.	.	.	.	.	.12	.13	.	.	.	.
6 vars	.79	.	.	.	.	-.06	.	.	.17	.	.	.	.	.07	.	.11	.10	.	.10	.	.
7 vars	.79	.	.	.	.	-.06	.	.	.17	.	.	-.05	.	.11	.	.12	.11	.	.10	.	.
8 vars	.80	-.08	.10	.	.	-.06	.	.	.15	.	.	.	.	.07	.	.12	.10	.	.09	.	.

Objective = 12 Class Size = Very Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.80	-.10	.10	.	.	-.05	.	.	.14	.	-.05	.	.	.12	.	.13	.10	.	.07	.04	.
3 vars	.79	.	.	.	.	.	.	.	.22	.	.	.	.	.	.	.15	.14	.	.	.	.
4 vars	.79	.	.	.	.	.	.	.	.18	.	.	.	.	.08	.	.15	.11	.	.	.	.
5 vars	.80	-.05	.	.	.	.	.	.	.18	.	.	.	.	.11	.	.16	.12	.	.	.	.
6 vars	.80	-.09	.10	.	.	.	.	.	.16	.	.	.	.	.09	.	.15	.11	.	.	.	.
7 vars	.80	-.09	.12	.	.	.	.	.	.16	.	-.04	.	.	.12	.	.15	.10	.	.	.	.
8 vars	.80	-.07	.	.	.	-.05	.	.	.17	.	.	.	.	.10	.	.13	.10	.	.09	.04	.

Appendix C: Bayesian Model Averaging on Overall Summary Measures

Excellence of teacher Class Size = Small

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.83	.14	.08		.02	-.01			.03	-.03	.24	-.01		.10	-.01	.03	-.04	.02			.01
3 vars	.82	.19									.29			.10							
4 vars	.82	.15	.09								.27			.07							
5 vars	.82	.15	.10								.26			.11			-.06				
6 vars	.82	.15	.09								.25			.11			-.06	.03			
7 vars	.83	.15	.08						.05	-.03	.26			.10			-.05				
8 vars	.83	.15	.08						.04	-.03	.25			.10			-.05	.02			

Excellence of teacher Class Size = Medium

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.87	.16	.08	.01	.01	.00			.02	-.02	.26	-.01		.10	-.01		-.05	.02	-.01	.02	
3 vars	.86	.19									.31			.08							
4 vars	.87	.20									.30			.13			-.06				
5 vars	.87	.16	.09								.28			.11			-.07				
6 vars	.87	.16	.08								.27			.11			-.07	.02			
7 vars	.87	.16	.09								.26			.11	-.02		-.06	.02			
8 vars	.87	.16	.08								.26			.11	-.02		-.07	.02		.02	

Excellence of teacher Class Size = Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.90	.13	.10					-.02	.02	-.02	.28	-.01	.02	.11			-.06	.02			
3 vars	.89	.14	.10								.35										
4 vars	.90	.19									.32			.13			-.07				
5 vars	.90	.13	.11								.30			.11			-.08				
6 vars	.90	.13	.10								.28		.03	.11			-.07				
7 vars	.90	.13	.11							-.02	.28		.03	.11			-.06				
8 vars	.90	.13	.10							-.02	.28		.02	.11			-.06	.01			

Excellence of teacher Class Size = Very Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.94	.12	.10		.04		-.06		.02	-.02	.32	-.02		.12			-.06	.02			
3 vars	.93	.16							.06		.38										
4 vars	.93	.18									.33			.12			-.06				
5 vars	.93	.13	.11								.31			.11			-.07				
6 vars	.93	.12	.10								.30			.10			-.07	.02			
7 vars	.94	.13	.10				-.05				.32			.13			-.07	.02			
8 vars	.94	.12	.11							-.02	.30	-.03		.12			-.06	.02			

Excellence of course. Class Size = Small

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.70	.02		.02	.07	-.03	.04	.02	.02	-.06	.13			.19	.03	.08	-.03		.02	.02	-.02
3 vars	.69										.20			.22		.10					
4 vars	.69				.10						.16			.18		.08					
5 vars	.70				.10					-.06	.16			.19		.11					
6 vars	.70				.10					-.05	.16			.21		.12	-.03				
7 vars	.70				.08		.04			-.05	.14			.20		.12	-.03				
8 vars	.70				.08		.04			-.06	.14			.19		.11	-.04			.03	

Excellence of course. Class Size = Medium

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.74	.01		.01	.09	-.04	.06		.01	-.06	.11	-.02	.01	.22	.04	.08	-.03	.01	.03	.02	-.02
3 vars	.73				.13						.15			.24							
4 vars	.73				.12						.15			.20		.07					
5 vars	.74				.12					-.06	.15			.21		.10					
6 vars	.74				.11					-.05	.14			.23		.11	-.04				
7 vars	.74				.09		.05			-.05	.12			.22		.11	-.04				
8 vars	.74				.09		.05			-.06	.13			.22	.02	.10	-.04				

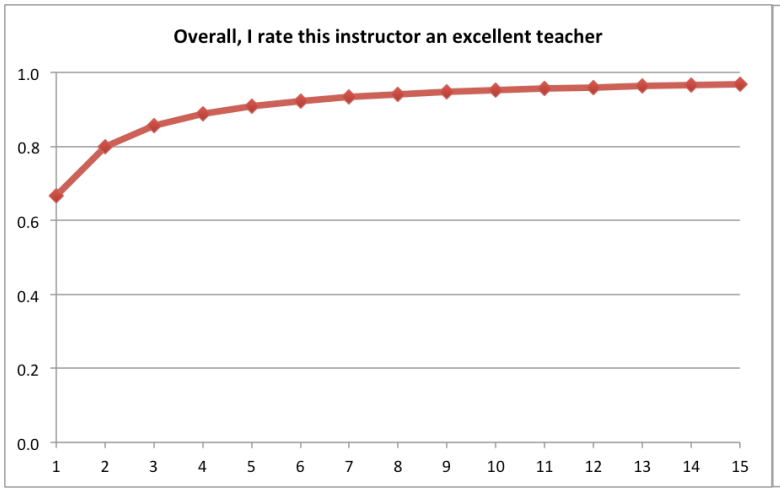
Excellence of course. Class Size = Large

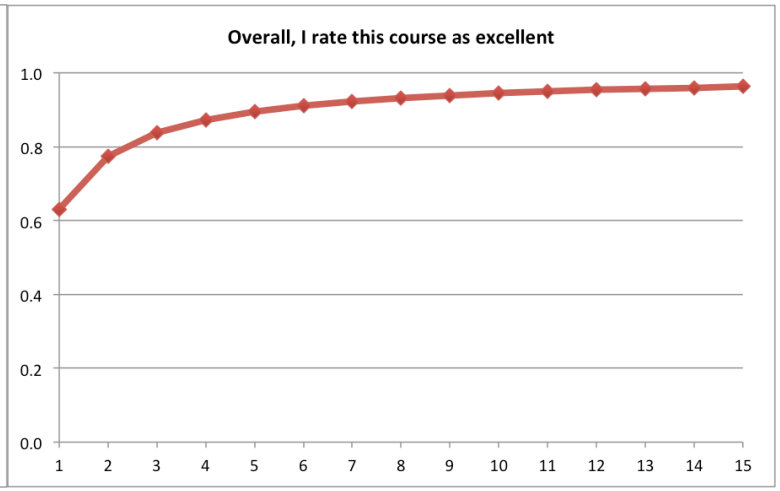
Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.78				.07	-.03	.04		.04	-.02	.13	-.02		.21	.02	.05		.01	.02		
3 vars	.77								.09		.19			.26							
4 vars	.77				.09						.16			.21		.08					
5 vars	.78				.08				.05		.16			.21		.05					
6 vars	.78				.08	-.01			.04		.15			.21		.06					
7 vars	.78				.08	-.01			.04		.14			.21		.06		.02			
8 vars	.78				.06	-.02	.04		.04		.13			.20		.06		.02			

Excellence of course. Class Size = Very Large

Model Size	R^2	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM17	TM18	TM19	TM20
Full	.80		-.07		.09	-.05			.04	-.03	.12			.25	.03	.11		.02	.06	-.03	
3 vars	.79										.15			.29		.11					
4 vars	.79										.14			.31		.15				-.06	
5 vars	.79				.10	-.04					.11			.24		.13					
6 vars	.80				.09						.09			.26		.13		.03		-.05	
7 vars	.80				.09	-.02					.09			.26		.14		.03		-.04	
8 vars	.80		-.06		.11	-.05				-.03	.12			.24		.14			.06		

Appendix D: Plots of Inter-class Reliability Coefficients for Overall Summary Measures





Appendix E: Inter-class Reliability Coefficients and Standard Errors of Measurement

	Class Number																													
	1		2		3		4		5		6		7		8		9		10		11		12		13		14		15	
Item	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*
Teaching methods																														
1. Displayed personal interest in students	.70	.25	.82	.19	.87	.16	.90	.14	.92	.13	.93	.12	.94	.11	.95	.10	.95	.10	.96	.09	.96	.09	.97	.09	.97	.08	.97	.08	.97	.08
2. Helped students answer own questions	.68	.28	.81	.22	.87	.18	.90	.16	.91	.15	.93	.13	.94	.13	.94	.12	.95	.11	.96	.11	.96	.10	.96	.10	.97	.09	.97	.09	.97	.09
3. Scheduled work helpfully	.63	.28	.78	.22	.84	.18	.87	.16	.90	.15	.91	.14	.92	.13	.93	.12	.94	.11	.95	.11	.95	.10	.95	.10	.96	.09	.96	.09	.96	.09
4. Demonstrated imp of subject	.67	.25	.81	.19	.86	.16	.89	.14	.91	.13	.93	.12	.94	.11	.94	.10	.95	.10	.95	.09	.96	.09	.96	.09	.96	.08	.97	.08	.97	.08
5. Formed teams, discussion groups	.65	.54	.79	.42	.85	.35	.88	.31	.90	.28	.92	.26	.93	.24	.94	.23	.94	.21	.95	.20	.95	.20	.96	.19	.96	.18	.96	.17	.97	.17
6. Made clear how topics fit	.65	.28	.79	.22	.85	.19	.88	.17	.90	.15	.92	.14	.93	.13	.94	.12	.94	.11	.95	.11	.95	.10	.96	.10	.96	.10	.96	.09	.97	.09
7. Explained criticisms	.68	.31	.81	.24	.86	.20	.89	.18	.91	.16	.93	.15	.94	.14	.94	.13	.95	.12	.95	.12	.96	.11	.96	.11	.96	.10	.97	.10	.97	.10
8. Stimulated intellectual effort	.65	.32	.79	.25	.85	.21	.88	.18	.90	.17	.92	.15	.93	.14	.94	.13	.94	.13	.95	.12	.95	.12	.96	.11	.96	.11	.96	.10	.97	.10
9. Encouraged use of multiple resources	.64	.37	.78	.29	.84	.24	.88	.22	.90	.20	.92	.18	.93	.17	.94	.16	.94	.15	.95	.14	.95	.14	.96	.13	.96	.13	.96	.12	.96	.12
10. Explained clearly	.69	.33	.81	.25	.87	.21	.90	.19	.92	.17	.93	.15	.94	.14	.95	.14	.95	.13	.96	.12	.96	.12	.96	.11	.97	.11	.97	.10	.97	.10
11. Related to real life	.68	.30	.81	.23	.87	.20	.90	.17	.92	.16	.93	.14	.94	.13	.95	.13	.95	.12	.96	.11	.96	.11	.96	.10	.97	.10	.97	.10	.97	.09
12. Tests covered important points	.58	.30	.74	.24	.81	.20	.85	.18	.87	.16	.89	.15	.91	.14	.92	.13	.93	.13	.93	.12	.94	.11	.94	.11	.95	.11	.95	.10	.95	.10
13. Introduced stimulating ideas	.67	.31	.80	.24	.86	.20	.89	.18	.91	.16	.92	.15	.93	.14	.94	.13	.95	.12	.95	.12	.96	.11	.96	.11	.96	.10	.97	.10	.97	.10
14. Involved students in hands on activities	.66	.43	.80	.34	.86	.28	.89	.25	.91	.23	.92	.21	.93	.19	.94	.18	.95	.17	.95	.16	.96	.16	.96	.15	.96	.15	.97	.14	.97	.14
15. Inspired students to set high goals	.68	.34	.81	.26	.86	.22	.89	.19	.91	.17	.93	.16	.94	.15	.94	.14	.95	.13	.95	.13	.96	.12	.96	.12	.96	.11	.97	.11	.97	.10
16. Asked students to share experiences	.75	.36	.86	.27	.90	.23	.92	.20	.94	.18	.95	.17	.96	.15	.96	.14	.97	.14	.97	.13	.97	.12	.97	.12	.98	.11	.98	.11	.98	.11
17. Provided timely feedback	.64	.33	.78	.26	.84	.22	.88	.20	.90	.18	.92	.16	.93	.15	.94	.14	.94	.13	.95	.13	.95	.12	.96	.12	.96	.11	.96	.11	.96	.11
18. Asked students to help each other	.66	.35	.79	.27	.85	.23	.89	.20	.91	.18	.92	.17	.93	.16	.94	.15	.95	.14	.95	.13	.95	.13	.96	.12	.96	.12	.96	.11	.97	.11
19. Assessments required creativity	.68	.34	.81	.26	.86	.22	.89	.20	.91	.18	.93	.16	.94	.15	.94	.14	.95	.13	.95	.13	.96	.12	.96	.12	.96	.11	.97	.11	.97	.11
20. Encouraged student/faculty contact	.67	.33	.80	.25	.86	.21	.89	.19	.91	.17	.92	.16	.93	.15	.94	.14	.95	.13	.95	.12	.96	.12	.96	.11	.96	.11	.97	.11	.97	.10
Learning objectives																														
21. Factual knowledge	.59	.30	.75	.24	.81	.20	.85	.18	.88	.16	.90	.15	.91	.14	.92	.13	.93	.12	.94	.12	.94	.11	.95	.11	.95	.10	.95	.10	.96	.10
22. Principles and theories	.60	.30	.75	.24	.82	.20	.86	.18	.88	.16	.90	.15	.91	.14	.92	.13	.93	.12	.94	.12	.94	.11	.95	.11	.95	.10	.95	.10	.96	.10
23. Applications	.62	.30	.77	.24	.83	.20	.87	.18	.89	.16	.91	.15	.92	.14	.93	.13	.94	.12	.94	.12	.95	.11	.95	.11	.96	.10	.96	.10	.96	.10

24. Professional skills, viewpoints	.62	.31	.77	.24	.83	.21	.87	.18	.89	.16	.91	.15	.92	.14	.93	.13	.94	.13	.94	.12	.95	.11	.95	.11	.96	.11	.96	.10	.96	.10
25. Team skills	.64	.44	.78	.34	.84	.29	.88	.26	.90	.23	.92	.21	.93	.20	.94	.19	.94	.18	.95	.17	.95	.16	.96	.16	.96	.15	.96	.14	.96	.14
26. Creative capacities	.76	.36	.86	.27	.90	.23	.93	.20	.94	.18	.95	.16	.96	.15	.96	.14	.97	.13	.97	.13	.97	.12	.97	.12	.98	.11	.98	.11	.98	.11
27. Broad liberal education	.76	.33	.86	.25	.90	.21	.93	.18	.94	.17	.95	.15	.96	.14	.96	.13	.97	.13	.97	.12	.97	.11	.97	.11	.98	.10	.98	.10	.98	.10
28. Communication skills	.73	.38	.85	.29	.89	.24	.92	.21	.93	.19	.94	.17	.95	.16	.96	.15	.96	.14	.96	.14	.97	.13	.97	.13	.97	.12	.97	.12	.98	.11
29. Find, use resources	.64	.34	.78	.26	.84	.22	.87	.20	.90	.18	.91	.17	.92	.15	.93	.14	.94	.14	.95	.13	.95	.12	.95	.12	.96	.12	.96	.11	.96	.11
30. Values development	.74	.33	.85	.25	.90	.21	.92	.18	.93	.17	.95	.15	.95	.14	.96	.13	.96	.13	.97	.12	.97	.11	.97	.11	.97	.11	.98	.10	.98	.10
31. Critical analysis	.70	.32	.82	.25	.87	.21	.90	.18	.92	.17	.93	.15	.94	.14	.95	.13	.95	.13	.96	.12	.96	.11	.97	.11	.97	.11	.97	.10	.97	.10
32. Interest in learning	.66	.31	.80	.24	.85	.21	.89	.18	.91	.16	.92	.15	.93	.14	.94	.13	.95	.12	.95	.12	.96	.11	.96	.11	.96	.10	.96	.10	.97	.10
Course ratings																														
33. Amount of reading	.68	.40	.81	.31	.86	.26	.89	.23	.91	.21	.93	.19	.94	.18	.94	.17	.95	.16	.95	.15	.96	.14	.96	.14	.96	.13	.97	.13	.97	.12
34. Amount of other work	.59	.35	.74	.28	.81	.24	.85	.21	.88	.19	.89	.18	.91	.17	.92	.16	.93	.15	.93	.14	.94	.13	.94	.13	.95	.12	.95	.12	.96	.12
35. Difficulty of subject matter	.64	.33	.78	.26	.84	.22	.88	.19	.90	.18	.91	.16	.93	.15	.93	.14	.94	.13	.95	.13	.95	.12	.96	.12	.96	.11	.96	.11	.96	.10
Self-ratings																														
36. Strong desire to take the course	.62	.41	.76	.32	.83	.27	.86	.24	.89	.22	.91	.20	.92	.19	.93	.18	.94	.17	.94	.16	.95	.15	.95	.15	.95	.14	.96	.14	.96	.13
37. Worked harder on this course than most	.57	.35	.73	.28	.80	.24	.84	.21	.87	.19	.89	.18	.90	.16	.91	.16	.92	.15	.93	.14	.94	.13	.94	.13	.95	.12	.95	.12	.95	.12
38. Wanted this instructor	.64	.40	.78	.31	.84	.26	.88	.23	.90	.21	.92	.19	.93	.18	.94	.17	.94	.16	.95	.15	.95	.15	.96	.14	.96	.14	.96	.13	.96	.13
39. Wanted course regardless of instructor	.61	.34	.76	.27	.82	.23	.86	.20	.89	.18	.90	.17	.92	.16	.93	.15	.93	.14	.94	.13	.95	.13	.95	.12	.95	.12	.96	.11	.96	.11
43. Usually work hard on academic work	.60	.21	.75	.16	.82	.14	.86	.12	.88	.11	.90	.10	.91	.10	.92	.09	.93	.09	.94	.08	.94	.08	.95	.08	.95	.07	.96	.07	.96	.07
Global ratings																														
40. Increase positive attitude toward field	.66	.34	.79	.26	.85	.22	.88	.20	.91	.18	.92	.16	.93	.15	.94	.14	.95	.14	.95	.13	.95	.12	.96	.12	.96	.11	.96	.11	.97	.11
41. Excellent instructor	.67	.35	.80	.27	.86	.23	.89	.20	.91	.18	.92	.17	.93	.15	.94	.15	.95	.14	.95	.13	.96	.13	.96	.12	.96	.12	.97	.11	.97	.11
42. Excellent course	.63	.36	.77	.28	.84	.24	.87	.21	.90	.19	.91	.18	.92	.16	.93	.15	.94	.15	.94	.14	.95	.13	.95	.13	.96	.12	.96	.12	.96	.11
Additional method items																														
44. Used variety of evaluation methods	.63	.33	.77	.26	.84	.22	.87	.20	.90	.18	.91	.16	.92	.15	.93	.14	.94	.14	.94	.13	.95	.12	.95	.12	.96	.11	.96	.11	.96	.11
45. Expected students to take responsibility	.63	.20	.77	.16	.83	.13	.87	.12	.89	.11	.91	.10	.92	.09	.93	.09	.94	.08	.94	.08	.95	.07	.95	.07	.96	.07	.96	.07	.96	.06
46. High achievement standards	.65	.24	.79	.19	.85	.16	.88	.14	.90	.13	.92	.12	.93	.11	.94	.10	.94	.10	.95	.09	.95	.09	.96	.08	.96	.08	.96	.08	.96	.07
47. Used educational technology	.72	.34	.84	.26	.89	.22	.91	.19	.93	.17	.94	.16	.95	.15	.95	.14	.96	.13	.96	.12	.97	.12	.97	.11	.97	.11	.97	.10	.98	.10

Appendix F: Inter-class Reliability Coefficients and Standard Errors of Measurement by Class Size

	Class size							
	10-14		15-34		35-49		50+	
Item	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*	r_{11}	*SEM*
Teaching methods								
1. Displayed personal interest in students	.63	.28	.70	.25	.77	.23	.78	.24
2. Helped students answer own questions	.61	.31	.69	.28	.74	.25	.77	.25
3. Scheduled work helpfully	.54	.33	.63	.27	.71	.24	.73	.25
4. Demonstrated imp of subject	.60	.29	.69	.24	.73	.23	.76	.23
5. Formed teams, discussion groups	.61	.54	.68	.51	.73	.50	.75	.48
6. Made clear how topics fit	.57	.33	.69	.27	.73	.25	.76	.24
7. Explained criticisms	.63	.35	.68	.31	.71	.28	.75	.27
8. Stimulated intellectual effort	.60	.36	.69	.30	.74	.26	.77	.26
9. Encouraged use of multiple resources	.61	.39	.68	.34	.72	.31	.73	.31
10. Explained clearly	.61	.37	.68	.33	.74	.30	.77	.29
11. Related to real life	.64	.33	.71	.29	.75	.26	.79	.25
12. Tests covered important points	.54	.35	.61	.30	.65	.25	.72	.25
13. Introduced stimulating ideas	.62	.34	.71	.29	.75	.27	.79	.26
14. Involved students in hands on activities	.65	.42	.70	.40	.74	.39	.77	.39
15. Inspired students to set high goals	.61	.37	.70	.31	.75	.29	.78	.28
16. Asked students to share experiences	.71	.40	.78	.34	.81	.32	.83	.32
17. Provided timely feedback	.58	.41	.65	.34	.68	.31	.71	.31
18. Asked students to help each other	.61	.37	.68	.33	.73	.30	.77	.29
19. Assessments required creativity	.66	.35	.71	.32	.73	.30	.75	.31
20. Encouraged student/faculty contact	.63	.36	.70	.31	.75	.28	.77	.26
Learning objectives								
21. Factual knowledge	.46	.37	.65	.28	.70	.25	.74	.23
22. Principles and theories	.51	.36	.66	.28	.69	.25	.73	.23
23. Applications	.55	.34	.66	.28	.72	.25	.76	.24
24. Professional skills, viewpoints	.51	.38	.68	.28	.73	.25	.76	.24

25. Team skills	.58	.49	.67	.42	.73	.39	.78	.35
26. Creative capacities	.71	.41	.77	.33	.77	.33	.80	.31
27. Broad liberal education	.69	.42	.77	.33	.80	.30	.84	.26
28. Communication skills	.67	.43	.77	.34	.76	.34	.79	.33
29. Find, use resources	.56	.41	.65	.33	.71	.29	.75	.27
30. Values development	.69	.40	.75	.32	.80	.28	.84	.26
31. Critical analysis	.63	.40	.73	.31	.77	.27	.82	.25
32. Interest in learning	.57	.38	.69	.30	.74	.27	.79	.24
Course ratings								
33. Amount of reading	.65	.46	.72	.38	.79	.29	.78	.31
34. Amount of other work	.53	.41	.60	.34	.71	.29	.79	.26
35. Difficulty of subject matter	.59	.38	.66	.32	.74	.28	.82	.25
Self-ratings								
36. Strong desire to take the course	.59	.42	.65	.39	.73	.32	.79	.27
37. Worked harder on this course than most	.52	.39	.62	.32	.71	.27	.80	.24
38. Wanted this instructor	.61	.44	.68	.38	.76	.31	.76	.31
39. Wanted course regardless of instructor	.56	.38	.63	.33	.70	.27	.81	.22
43. Usually work hard on academic work	.52	.27	.71	.18	.70	.16	.66	.16
Global ratings								
40. Increase positive attitude toward field	.58	.38	.69	.32	.75	.28	.77	.27
41. Excellent instructor	.59	.40	.68	.34	.73	.31	.77	.31
42. Excellent course	.55	.40	.66	.34	.73	.30	.76	.29
Additional method items								
44. Used variety of evaluation methods	.55	.37	.66	.31	.72	.30	.73	.32
45. Expected students to take responsibility	.54	.26	.68	.20	.71	.17	.75	.16
46. High achievement standards	.56	.30	.70	.22	.75	.19	.78	.18
47. Used educational technology	.61	.43	.74	.33	.77	.30	.81	.25

Appendix G: Expert Panels

Expert Panel 1

Laura Brannon
Debra Fowler
Tim Frey
Gerald Hanna
Steve Horvath
Roger McHaney
Todd McLoda
Loraine Phillips
Barbara Plake
Michael Stankey
Tracee Synco
Marilla Svinicki
Tiffany van der Merwe
Jon Wergin
Karl Wirth

Expert Panel 2

Larry Braskamp
Robert Colvin
Stephanie Juillerat
Andrea Karkowski
Claudine Keenan
Kristin Keyes
Colleen Pilgrim
Wayne Pricer
Richard J. Sherry
Peter R. Skoner
David Starrett
Alec Thomson
Susan Wooten

Appendix H: SAS Code for Reliability Analyses

Code for Proc HPMIXED

```
proc hpmixed data  =  output.combined method = reml noclprint;
      class faculty_id course_id;
      model res_&q  =  ;
      random int/subject  =  faculty_id;
      random int/subject  =  course_id(faculty_id);
      ods output covparms  =  output.ests;
```

Followed by:

```
data int1;
      set output.ests;
      if covparm  =  "Intercept" and subject = "faculty_id";

      cov_faculty_int  =  estimate;
      item  =  "&q  ";

data int2;
      set output.ests;
      if covparm  =  "Intercept" and
                  subject  =  "course_id(faculty_i)";

      cov_course_int  =  estimate;
      item  =  "&q  ";

data res;
      set output.ests;
      if covparm  =  "Residual";

      cov_res  =  estimate;
      item  =  "&q  ";

data rel_&q;
      merge int1 int2 res;
      reliability_faculty  =  cov_faculty_int/
                              (cov_res+cov_faculty_int);
      reliability_course  =  cov_course_int/
                              (cov_res+cov_course_int);
      keep item reliability_faculty reliability_course;

proc print data  =  rel_&q;
```

Appendix I: Item by Item Comparisons of the Original and Proposed Updates of Learning Objectives

Current Learning Outcomes	New Learning Outcomes	Differences between existing and new items	Rationale for change[4]
1. (21*) Gaining factual knowledge (terminology, classifications, methods, trends)	1. Gaining a basic understanding of the subject (e.g., factual knowledge, methods, principles, generalizations, theories)	Significant change: Merged the first two objectives into a single learning outcome	Faculty ratings of importance on existing Objectives 21 and 22 are highly correlated. Faculty who select one of those two objectives tend to also select the other. In turn, student ratings of progress on those objectives are highly correlated. So, there is considerable redundancy.
2. (22) Learning fundamental principles, generalizations, or theories		Removed: Merged with Objective 1	See above
	2. Developing knowledge and understanding of diverse perspectives, global awareness, or other cultures	New item: Created learning outcome to fill gap related to diversity and global awareness	This new learning outcomes addresses AAC&U VALUE rubrics "Intercultural Knowledge and Competence" and "Global Learning."
3. (23) Learning to *apply* course material (to improve thinking, problem solving, and decisions)	3. Learning to *apply* course material (to improve thinking, problem solving, and decisions)	No change	
4. (24) Developing specific skills, competencies, and points of view needed by professionals in the field most closely related to this course	4. Developing specific skills, competencies, and points of view needed by professionals in the field most closely related to this course	No change	
5. (25) Acquiring skills in working with others as a member of a team	5. Acquiring skills in working with others as a member of a team	No change	
6. (26) Developing creative capacities (writing, inventing, designing, performing in art, music, drama, etc.)	6. Developing creative capacities (inventing; designing; writing; performing in art, music, drama, etc.)	Minor revision (no change to meaning): reordered parenthetic examples	This change was an attempt to address the concern expressed by some students and faculty that the current Objective 26 pertains only to writing and art courses. The order of examples was rearranged to

[4] Changes are based on extensive research and feedback from expert panels and focus groups

* Numbers in parentheses reference numbers as they currently appear on the Diagnostic Feedback instrument

			highlight “inventing” and “designing.”
7. (27) Gaining a broader understanding and appreciation of intellectual/cultural activity (music, science, literature, etc.)	7. Gaining a broader understanding and appreciation of intellectual/cultural activity (music, science, literature, etc.)	No change	
8. (28) Developing skill in expressing myself orally or in writing	8. Developing skill in expressing myself orally or in writing	No change	
9. (29) Learning how to find and use resources for answering questions or solving problems	9. Learning how to find, evaluate, and use resources to explore a topic in depth	Significant change: Blends information literacy (current Objective 29) and lifelong learning (current Objective 32) outcomes into one objective	This change was made in response to AAC&U VALUE rubric “Information Literacy,” which places an emphasis on *evaluating* resources. The phrase “explore a topic in depth” was included to address the “Skills for Lifelong Learning” VALUE rubric.
10. (30) Developing a clearer understanding of, and commitment to, personal values	10. Developing ethical reasoning and/or ethical decision making	New item: Replaces existing “personal values” Objective 10 with ethical reasoning	This outcome was added to address AAC&U VALUE rubric “Ethical Reasoning.”
11. (31) Learning to analyze and critically evaluate ideas, arguments, and points of view	11. Learning to *analyze* and *critically* evaluate ideas, arguments, and points of view.	No change	
12. (32) Acquiring an interest in learning more by asking my own questions and seeking answers		Removed existing item related to independent learning	
	12. Learning to apply knowledge and skills to benefit others or serve the public good.	New item: Created new learning outcome to address civic engagement	This outcome was added to address AAC&U VALUE rubric “Civic Engagement.”
	13. Learning appropriate methods for collecting, analyzing, and interpreting numerical information	New item: Created new learning outcome to address quantitative literacy	This outcome was added to address AAC&U VALUE rubric “Quantitative Literacy.”

Appendix J: Item by Item Comparisons of the Original and Proposed Updates of Teaching Methods

Current Teaching Methods	New Teaching Methods	Differences between existing and new items	Rationale for changes[5]
1. Displayed a personal interest in students and their learning	1. Displayed a personal interest in students and their learning	No change	
2. Found ways to help students answer their own questions	2. Found ways to help students answer their own questions	No change	
3. Scheduled course work (class activities, tests, projects) in ways that encouraged students to stay up- to-date in their work		Removed	This method is not a significant variable in any of the regression models for predicting progress on relevant learning objectives or ratings of overall summary measures.
	3.Helped students to interpret subject matter from diverse perspectives (e.g., different cultures, religions, genders, political views)	New item: Added teaching method related to diversity	Focus groups, expert panels, IDEA Updating Team, and faculty recommended it. Supported by literature on teaching and learning.
4.Demonstrated the importance and significance of the subject matter	4.Demonstrated the importance and significance of the subject matter	No change	
5.Formed "teams" or "discussion groups" to facilitate learning	5.Formed teams or groups to facilitate learning	Minor revision (no change to meaning): Removed "discussion"	"Discussion" groups is too limiting. There are many different types of and purposes of groups.
6.Made it clear how each topic fit into the course		Removed	Conceptually similar to #4 and #10. When we removed #6 from regression models, there was no significant drop in variance explained.
	6.Encouraged students to reflect on and evaluate what they have learned	New item: Created teaching method that captured reflection and critical thinking	Focus groups, expert panels, IDEA Updating Team, and faculty recommended it.

[5] Changes are based on extensive research and feedback from experts panels and focus groups

			Supported by literature on teaching and learning.
7.Explained the reasons for criticisms of students' academic performance	7.Provided meaningful feedback on students' academic performance	Revised: Combined current Teaching Method 7 with current Teaching Method 17	"Criticisms" conveyed negative connotation.
8.Stimulated students to intellectual effort beyond that required by most courses	8.Stimulated students to intellectual effort beyond that required by most courses	No change	
9.Encouraged students to use multiple resources (e.g. data banks, library holdings, outside experts) to improve understanding	9.Encouraged students to use multiple resources (e.g. Internet, library holdings, outside experts) to improve understanding	Minor revision (no change to meaning): Replaced "databanks" with "Internet"	Language in current Teaching Method 9 was outdated.
10.Explained course material clearly and concisely	10.Explained course material clearly and concisely	No change	
11.Related course material to real life situations	11.Related course material to real life situations	No change	
12.Gave tests, projects, etc. that covered the most important points of the course	12.Gave tests, projects, etc. that covered the most important points of the course	No change	
13.Introduced stimulating ideas about the subject	13.Introduced stimulating ideas about the subject	No change	
14. Involved students in "hands on" projects such as research, case studies, or "real life" activities	14. Involved students in hands-on projects such as research, case studies, or real life activities	Dropped unnecessary quotation marks and added hyphen	
15. Inspired students to set and achieve goals which really challenged them	15. Inspired students to set and achieve goals which really challenged them	No change	
16. Asked students to share ideas and experiences with others whose backgrounds and viewpoints differ from their own	16. Asked students to share ideas and experiences with others whose backgrounds and viewpoints differ from their own	No change	

17. Provided timely and frequent feedback on tests, reports, projects, etc. to help students improve		Removed: Combined with Teaching Method 7	This method is not a significant variable in any of the regression models for predicting progress on relevant learning objectives or ratings of overall summary measures.
	17. Created opportunities for students to apply course content outside the classroom.	New: Created to capture use of community service as a teaching method	Focus groups, expert panels, IDEA Updating Team, and faculty recommended it. Supported by literature on teaching and learning.
18. Asked students to help each other understand ideas or concepts	18. Asked students to help each other understand ideas or concepts	No change	
19. Gave projects, tests, or assignments that required original or creative thinking	19. Gave projects, tests, or assignments that required original or creative thinking	No change	
20. Encouraged student-faculty interaction outside of class (office visits, phone calls, e-mail, etc.)		Removed	This method is not a significant variable in any of the regression models for predicting relevant learning objectives or ratings of overall summary measures.