



IDEA Technical  
Report No. 19

**Analysis of IDEA Student Ratings of  
Instruction System 2015 Pilot Data**

Dan Li  
Stephen L. Benton  
Ron Brown  
Patricia Sullivan  
Kenneth R. Ryalls

## Table of Contents

Executive Summary.....	6
Introduction .....	7
Why Was a Revision Necessary?.....	7
Overview of Changes in IDEA SRI <i>Diagnostic Feedback</i> 2016 .....	8
Changes in Teaching Methods .....	8
Changes in Learning Objectives.....	9
Changes in Student and Course Characteristics .....	9
Changes in Variables used in the Adjusted Score Models .....	10
Preparing for the Pilot Study .....	10
Pilot Learning Objectives .....	10
Pilot Teaching Methods.....	10
Pilot Course and Student Characteristics .....	11
Procedures Involved in Setting up the Pilot Study.....	11
Pilot Sample.....	11
Comparison of Student Ratings in the Pilot Sample and 2014-2015 IDEA Research Dataset .....	12
Analyses of the Pilot Research Dataset .....	14
Comparison of Faculty Ratings of Learning Objectives in the Pilot Sample and 2014-2015 Research Dataset .....	15
Selecting Objectives for the Revised Instruments.....	17
Correlations Between Student and Faculty Ratings of Existing and Proposed Learning Objectives .....	18
Correlations Among Faculty Ratings of Relevance of Learning Objectives .....	21
Factor Structure of Faculty Ratings of Relevance on Learning Objectives .....	23
Factor Structure of Student Ratings of Progress on Learning Objectives.....	24
Student Progress on Relevant Objectives.....	26
Selecting Teaching Methods for IDEA2 <i>Diagnostic Feedback</i> .....	28
Relationships between Teaching Methods and Relevant Learning Objectives.....	28
Factor Structure of Student Ratings of Teaching Methods .....	31
Overall Summary Measures .....	33
Selecting Variables for the Adjusted Score Formulas.....	34
Course Characteristics .....	34
Student Characteristics .....	35
Residual Difficulty and Effort Scores .....	35
Background Preparation and Self-Efficacy .....	36
Reliabilities .....	40
Updated Student Ratings of Instruction Instruments.....	42
Diagnostic Feedback .....	42
Teaching Methods .....	42
Learning Objectives.....	43
Course Characteristics .....	44
Student Characteristics .....	44

Deleted: 25

Deleted: 27

Deleted: 29

Deleted: 29

Deleted: 34

Deleted: 43

Deleted: 43

Deleted: 43

	3
Overall Summary Measures .....	44
Learning Essentials.....	44
Learning Objectives.....	44
Student Characteristics .....	44
Overall Summary Measures .....	45
Suggested Actions for Teaching Methods.....	45
Appendix A: Item by Item Comparisons of the Original and Proposed Updates of Learning Objectives .....	51
Appendix B: Item by Item Comparisons of the Original and Proposed Updates of Teaching Methods .....	54
Appendix C: IDEA Alignment with NSSE Engagement Indicators .....	57
Appendix D: Bayesian Model Averaging on 13 Learning Objectives in Medium Classes (15-34) .....	60

Deleted: 45

## List of Tables

Table 1 Frequency of Levels of Carnegie Classification and Institutional Control by Regional Locations for Pilot Sample Institutions ( $N = 27$ ).....	12	Deleted: 11
Table 2 Means and Standard Deviations for Student Ratings of Existing Items on IDEA Diagnostic Feedback .....	13	Deleted: 12
Table 3 Means and Standard Deviations for Student Ratings of Pilot Items on IDEA Diagnostic Feedback ( $n = 3,484$ ).....	15	Deleted: 14
Table 4 Descriptive Statistics for Faculty Ratings of the Relevance of 12 Existing Learning Objectives in the 2014-2015 Research Dataset.....	16	Deleted: 15
Table 5 Descriptive Statistics for Faculty Ratings of the Importance of Existing and Proposed Learning Objectives in the Pilot Sample .....	17	Deleted: 16
Table 6 Correlations for Faculty Ratings of Relevance and Student Ratings of Progress on 12 Existing and 6 Proposed Learning Objectives.....	19	Deleted: 18
Table 7 Frequency of Instructor Ratings of Relevance on Learning Objectives Factual Knowledge, Principles and Theories, and Understanding Subject Matter in the Pilot Sample ....	20	Deleted: 19
Table 8 Correlations for Faculty Ratings of Relevance on 12 Existing and 6 Proposed Learning Objectives .....	22	Deleted: 21
Table 9 Factor Loadings from Principal Components Factor Analysis With Varimax Rotation: Eigenvalues and Percentages of Variance for Faculty Ratings of Importance on 12 Existing and 6 Proposed Learning Objectives.....	23	Deleted: 22
Table 10 Factor Loadings from Principal Components Factor Analysis With Varimax Rotation: Eigenvalues and Percentages of Variance for Faculty Ratings of Importance on 13 Learning Objectives in Diagnostic Feedback 2016 .....	24	Deleted: 23
Table 11 Factor Loadings from Principal Components Factor Analysis With Varimax Rotation: Eigenvalues and Percentages of Variance for Student Ratings of Progress on 12 Existing and 6 Proposed Learning Objectives.....	25	Deleted: 24
Table 12 Factor Loadings from Principal Components Factor Analysis With Varimax Rotation: Eigenvalues and Percentages of Variance for Student Ratings of Progress on Learning Objectives on IDEA2.....	26	Deleted: 25
Table 13 Means and Standard Deviations for Student Ratings on Relevant Learning Objectives in the Pilot Sample.....	27	Deleted: 26
Table 14 Teaching Methods Significantly Related with Progress on Learning Objectives in Medium-sized Classes (15 - 34) Methods Significantly Related with Progress on Learning Objectives in Medium-sized Classes (15 - 34).....	30	Deleted: 29
Table 15 Factor Loadings from Principal Components Factor Analysis With Varimax Rotation: Eigenvalues and Percentages of Variance for Student Ratings of 17 Existing and 4 Proposed Teaching Methods .....	31	Deleted: 30
Table 16 Factor Loadings from a Principal Components Factor Analysis with Varimax Rotation: Eigenvalues and Percentages of Variance for Student Ratings of Teaching Methods on IDEA2		

Diagnostic Feedback .....	33	Deleted: 32
Table 17 Hierarchical Regression Analysis Summary for Workload Variables Predicting Difficulty of Subject Matter ( $N = 3,484$ ).....	34	
Table 18 Correlations for Adjusted Score Model Independent Variables.....	36	Deleted: 35
Table 19 Regression Analyses Predicting Student Ratings on Relevant Learning Objectives and Summary Measures with Course Motivation, Work Habits, Enrollment, Difficulty Residual, Effort Residual, Background Preparation, and Self-efficacy .....	37	Deleted: 36
Table 20 Regression Analyses Predicting Student Ratings on Relevant Learning Objectives and Summary Measures with Course Motivation, Work Habits, Enrollment, Difficulty Residual, and Background Preparation .....	38	Deleted: 37
Table 21 Regression Analyses Predicting Student Ratings on Relevant Learning Objectives and Summary Measures with Course Motivation, Work Habits, Enrollment, and Background Preparation.....	39	Deleted: 38
Table 22 Within-group Interrater Reliability Coefficients and Standard Errors of Measurement of Items on the Pilot Diagnostic Feedback Instrument.....	41	Deleted: 39
Table 23 Average Student Ratings for Existing Teaching Method Items in Diagnostic Feedback by Class Size and Level of Course Motivation .....	45	Deleted: 44
Table 24 Average Student Ratings for Proposed Teaching Methods by Class Size and Average Student Ratings for Proposed Teaching Methods by Class Size and Level of Course Motivation in Medium-Sized Classes .....	48	Deleted: 46
Table 25 Estimated Average Student Ratings for Proposed Teaching Method Items in Diagnostic Feedback by Class Size and Level of Course Motivation .....	48	Deleted: 47

### Executive Summary

This report describes statistical analyses performed on data collected in the Spring of 2015 from the pilot study of proposed revised and new items in the IDEA Student Ratings of Instruction (SRI) system. What follows is a description of the methods employed, results obtained, and decisions made in selecting items for the updated instruments. The procedures explained herein occurred across a one-year period, beginning in the Spring of 2015 and ending with the development of IDEA *Diagnostic Feedback* (DF) 2016 and *Learning Essentials* (LE) 2016 in the Fall of 2015.

The need for the current revision stemmed from many changes that have occurred in higher education: the increase in the number of courses offered online; widespread agreement about global student learning outcomes not emphasized in 1999, notably civic engagement, ethical reasoning, diverse perspectives, and quantitative literacy; teaching methods that provide meaningful feedback and help students to interpret subject matter from diverse perspectives, encourage self-reflection, and engage students in service; student characteristics such as background preparation and self-efficacy that are known to influence learning; and the prevalence of mobile devices, which makes in-class capture more manageable but also makes a shorter instrument more desirable.

We turned to multiple information sources to guide our decision making: extensive statistical analyses of the Spring 2015 pilot dataset, analyses of aggregated ratings in the IDEA SRI 2011 to 2015 research datasets, the professional literature on teaching and learning, the Association of American Colleges and Universities (AAC&U) VALUE rubrics, and the National Survey of Student Engagement (NSSE) indicators. We also consulted with measurement experts, statisticians, the IDEA Updating Team, which included former higher-education administrators and faculty, as well as panels of experts in faculty development and evaluation.

The descriptive statistics found in the Spring 2015 pilot classes ( $N = 3,484$ ) were very similar to those in the IDEA SRI 2014 to 2015 research dataset, boosting confidence in the integrity of the pilot sample. Additional analyses led to new recommendations for specific teaching methods associated with student progress on relevant learning objectives. In addition, subscales were created for Instructor-Centered and Student-Centered teaching methods. Reliability coefficients for both subscales and all individual items were high.

The resulting 40-item *Diagnostic Feedback* 2016 includes 19 teaching methods, 13 learning objectives, 6 student and course characteristics, and 2 summary items. Three existing teaching methods were removed and three new ones added to assess diverse perspectives, student self-reflection, and service learning. Four new learning objectives were added to measure diverse perspectives, civic engagement, quantitative literacy, and ethical reasoning. Others were either dropped or modified to better align with widely accepted global learning outcomes. Student characteristics now include two new items: background preparation and self-efficacy. Finally, changes were made to the adjusted score formulas used to control for variables that are not under control of the instructor but can affect student ratings.

## Introduction

This report describes statistical analyses performed on data collected in the Spring of 2015 from the pilot study of proposed revised and new items in the IDEA Student Ratings of Instruction (SRI) system, *Diagnostic Feedback* (DF) 2016. What follows is a description of the methods employed, results obtained, and decisions made in selecting items for the updated instruments. Technical Report No. 18, *Revising the IDEA Student Ratings of Instruction System* (Benton, Li, Brown, Guo, & Sullivan, 2015), describes the processes involved in developing the pilot instrument. The procedures explained herein occurred across a one-year period, beginning in the Spring of 2015 and ending with the development of IDEA *Diagnostic Feedback* (DF) 2016 and *Learning Essentials* (LE) 2016 in the Fall of 2015.

We turned to multiple information sources to guide our decision making: extensive statistical analyses of the Spring 2015 pilot dataset, analyses of aggregated ratings in the IDEA SRI 2011 to 2015 research datasets, the professional literature on teaching and learning, the Association of American Colleges and Universities (AAC&U) VALUE rubrics, and the National Survey of Student Engagement (NSSE) indicators. We also consulted with measurement experts, statisticians, and the IDEA Updating Team, which included former higher-education administrators and faculty, as well as experts in faculty development and evaluation.

### Why Was a Revision Necessary?

Donald P. Hoyt (1973a) developed the original IDEA SRI system in 1969 with the help of faculty and students. Over the next two years, he made slight modifications that remained intact until 1975 when, with the support of a grant from the Kellogg Foundation, Hoyt and colleagues developed a version that remained largely unchanged for more than two decades (see Hoyt & Cashin, 1977, for a description of the process). Then, because of widespread changes in instructional approaches, essential learning objectives, and the settings in which instruction occurs, staff at The IDEA Center made more modifications in 1999 (Hoyt, Chen, Pallett, & Gross, 1999). That 47-item instrument remains today in paper format only. Since 1975, 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 stemmed from many changes that have occurred in higher education. The increase in the number of courses offered online is but one example. There is also widespread agreement among influential higher education organizations about global student learning outcomes, some of which were not emphasized in 1999, notably civic engagement, ethical reasoning, diverse perspectives and global awareness, and quantitative literacy. Teaching methods that provide meaningful feedback and help students to interpret subject matter from diverse perspectives, encourage self-reflection and self-evaluation, and engage students in service are now receiving more emphasis. Increasing understanding of teaching and learning also calls for a revisit of the long-standing IDEA instrument. Student characteristics such as background preparation and self-efficacy are known to influence learning, although only the former was included in the existing IDEA SRI system. The prevalence of

mobile devices makes in-class capture more manageable, though at the same time makes a shorter instrument more desirable.

In light of these changes, we embarked on an update of IDEA SRI that would incorporate contemporary learning outcomes and teaching methods, control for important student characteristics, and include only items that contribute greatly to summative and formative feedback. 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 beyond the instructor's control
- A focus on instructional improvement

#### **Overview of Changes in IDEA SRI *Diagnostic Feedback* 2016**

The 40-item *Diagnostic Feedback* 2016 includes 19 teaching methods, 13 learning objectives, 6 student and course characteristics, and 2 summary items. Appendices A and B provide an explanation of the similarities and differences between the previous and updated instruments. A brief summary of the changes follows below.

##### *Changes in Teaching Methods*

Newly added teaching methods include:

- Diverse perspectives: “Helped students to interpret subject matter from diverse perspectives (e.g., different cultures, religions, genders, political views)”
- Self-reflection: “Encouraged students to reflect on and evaluate what they have learned”
- Service learning: “Created opportunities for students to apply course content outside the classroom”

In addition, two existing items—“Provided timely and frequent feedback on tests, reports, projects, etc. to help students improve,” and “Explained the reasons for criticisms of students’ academic performance”—were synthesized into a new item, “Provided meaningful feedback on students’ academic performance.”

The following teaching methods were removed because they were not highly correlated with student progress on any of the learning objectives:

- “Displayed a personal interest in students and their learning”
- “Scheduled course work (class activities, tests, projects) in ways that encouraged students to stay up-to-date in their work”
- “Gave tests, projects, etc. that covered the most important points of the course”

Finally, the item “Formed ‘teams’ or ‘discussion groups’ to facilitate learning” was modified to “Formed teams or groups to facilitate learning.” Focus group participants and expert panel members believed the quotation marks were unnecessary. The word *discussion* was removed because collaborative learning encompasses various forms of acquiring knowledge not limited to discussion.

### *Changes in Learning Objectives*

The following changes affect IDEA *Diagnostic Feedback* and *Learning Essentials*. Six new learning objectives were added:

- Diverse perspectives: “Developing knowledge and understanding of diverse perspectives, global awareness, or other cultures”
- Civic engagement: “Learning to apply knowledge and skills to benefit others or serve the public good”
- Quantitative literacy: “Learning appropriate methods for collecting, analyzing, and interpreting numerical information”
- Ethical reasoning: “Developing ethical reasoning and/or ethical decision making,” which replaces “Developing a clearer understanding of, and commitment to, personal values”
- Information literacy: “Learning how to find, evaluate, and use resources to explore a topic in depth,” which replaces “Learning how to find and use resources for answering questions or solving problems” and “Acquiring an interest in learning more by asking my own questions and seeking answers”
- Understanding subject matter: “Gaining factual knowledge (terminology, classifications, methods, trends), which replaces “Learning fundamental principles, generalizations, or theories” and “Gaining a basic understanding of the subject (e.g., factual knowledge, methods, principles, generalizations, theories)”

### *Changes in Student and Course Characteristics*

Two existing items, “Amount of reading” and “Amount of work in other (non-reading) assignments,” were synthesized into a single item, “Amount of coursework.” Further, “My background prepared me well for this course’s requirements,” an item on the *Learning Essentials* instrument (previously named the *Short Form*), has also been added along with a new item on student self-efficacy: “When this course began I believed I could master its content.”

Three items measuring student characteristics were removed, because they were less important than other items in either predicting student progress on relevant learning objectives or computing adjusted scores:

- “I had a strong desire to take this course.”
- “I worked harder on this course than on most courses I have taken.”
- “I really wanted to take a course from this instructor.”

The following four method items were removed due to their experimental nature:

- “The instructor used a variety of methods—not only tests—to evaluate student progress on course objectives.”
- “The instructor expected students to take their share of responsibility for learning.”
- “The instructor had high achievement standards in this class.”
- “The instructor used educational technology (e.g. Internet, email, computer exercises, multi-media presentations) to promote learning.”

Lastly, we removed one of the summary items, “As a result of taking this course, I have

more positive feelings toward this field of study,” because it produces ambiguous information. 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.

#### *Changes in Variables used in the Adjusted Score Models*

Mean scores on several items pertaining to student and course characteristics are used to compute adjusted scores on student ratings of progress on relevant learning objectives and overall ratings of the instructor and course. The adjustments control for extraneous factors that can affect ratings but are beyond the instructor’s control. They are intended to level the playing field between instructors teaching students who vary in motivation and work habits and classes that differ in enrollment size and subject-matter difficulty. *Course motivation* (“I really wanted to take this course regardless of who taught it”), *work habits* (“As a rule, I put forth more effort than other students on academic work”), *class size*, and a residual of *subject matter difficulty* (“Difficulty of subject matter”) remain in the adjusted score models for the DF. *Background preparation* (“My background prepared me well for this course’s requirements”) has been added; whereas a residual of *student effort* (“I worked harder on this course than on most courses I have taken”) no longer plays a role in the calculations. *Learning Essentials* adjusted score models include *course motivation*, *work habits*, *class size*, and *background preparation*.

#### **Preparing for the Pilot Study**

In 2012, IDEA began what would be a four-year process to update the SRI instruments, culminating in a successful pilot of 13 new and revised items in Spring 2015. Based on procedures described in Benton et al. (2015), decisions were made to include the following pilot items, numbered 49 to 61, at the end of the existing 47-item DF instrument on the platform powered by Campus Labs. In the *Objective Selection Form* (previously called the *Faculty Information Form*), the six pilot learning objectives were displayed below the 12 original learning objectives as Items 13 to 18. The instructions and response options for the pilot items were the same as for existing items.

#### *Pilot Learning Objectives*

- “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”
- “Learning how to find, evaluate, and use resources to explore a topic in depth”
- “Gaining a basic understanding of the subject (e.g., factual knowledge, methods, principles, generalizations, theories)”

#### *Pilot Teaching Methods*

- “Helped students to interpret subject matter from diverse perspectives (e.g., different cultures, religions, genders, political views)”
- “Encouraged students to reflect on and evaluate what they have learned”
- “Created opportunities for students to apply course content outside the classroom”
- “Provided meaningful feedback on students’ academic performance”

### *Pilot Course and Student Characteristics*

- “Amount of coursework”
- “My background prepared me well for this course’s requirements”
- “When this course began I believed I could master its content”

### *Procedures Involved in Setting up the Pilot Study*

All institutions using the IDEA SRI on the Campus Labs platform were automatically included in the pilot study, although any institution could choose not to participate. Throughout the fall and spring terms, IDEA employed a multi-faceted communication plan to inform institutions of the pilot process. A webpage was created with supporting documents, including instrument crosswalks, rationale for changes, and sample instruments.

*Student Experience.* It was important to maintain the integrity of the existing SRI instrument; therefore, rather than integrating pilot items into the existing survey (i.e., listing new learning objectives along with the existing learning objectives), the pilot questions were displayed at the *end of the survey* as “additional questions.” The student experience, therefore, was unchanged when compared with previous administrations regarding items used to generate faculty reports. The piloted items were in the same format as the rest of the survey. Instructions for the new learning objectives told students to describe the amount of progress made on each outcome “even if the wording is similar to previously asked survey items.”

*Faculty Experience.* The new learning objectives were added to the *Objectives Selection Form (OSF)*. They were placed in a separate section following the original 12 learning objectives used to generate the individual faculty report. The piloted items were clearly identified with a banner “Pilot Learning Objectives.” The following statement was included on the OSF as well as a direct link to the IDEA webpage that provided additional information about the pilot process.

*Your response to these items will NOT affect your current student ratings. However, your thoughtful response will contribute to the revision of the Diagnostic Feedback instrument, which will provide feedback on your teaching in the future.*

### *Pilot Sample*

In the spring of 2015, 49,803 distinct students completed 128,600 surveys in 14,521 courses on the Campus Labs platform. The pilot sample was created by excluding courses that had an ineligible survey start date ( $n = 2,364$ ), where the instructor identified no existing objective as *Essential* or *Important* ( $n = 3,045$ ), and those with fewer than five student responses ( $n = 2,707$ ). As a result, 6,405 classes from 27 institutions, representing all regions of the continental U.S., were included in the pilot sample. Table 1 presents a breakdown of institutions in the pilot sample, based on their Higher Learning Commission (HLC) accreditation regions by Carnegie classification and by institutional control (public vs. private). Among the 27 institutions, 70% were public, and Baccalaureate and Master’s degree granting institutions each comprised 41% of the pilot sample institutions.

Table 1

*Frequency of Levels of Carnegie Classification and Institutional Control by Regional Locations for Pilot Sample Institutions (N = 27)*

Regional accreditation	Carnegie classification				Institutional control	
	Associate	Baccalaureate	Masters	Doctoral	Private	Public
Middle States	0	1	3	0	0	4
New England	0	0	1	0	0	1
North Central	2	6	3	1	4	8
Northwest	0	1	0	0	0	1
Southern	1	3	3	1	4	4
Western	0	0	1	0	0	1

*Comparison of Student Ratings in the Pilot Sample and 2014-2015 IDEA Research Dataset*

Prior to embarking upon an analysis of the pilot data we examined the degree to which ratings in the pilot sample were similar to those in the most recent annual IDEA research dataset. Table 2 presents means and standard deviations for the 47 items in the original DF instrument from the academic year 2014-2015 and the same items from the 2015 pilot (separating classes with a minimum of 5 and 10 responses). For all close-ended items, students responded using 5-point scales. To report how frequently students perceived that an instructor used each of the 20 teaching methods (Items 1 to 20) they responded, 1 = *Hardly Ever*, 2 = *Occasionally*, 3 = *Sometimes*, 4 = *Frequently*, 5 = *Almost Always*; for reporting their own progress on each of 12 learning objectives (Items 21 to 32), 1 = *No apparent progress*; 2 = *Slight progress*; 3 = *I made small gains on this objective*; 4 = *Moderate progress*; 5 = *I made some gains on this objective*; 6 = *Substantial progress*; 7 = *I made large gains on this objective*; 8 = *Exceptional progress*; 9 = *I made outstanding gains on this objective*; and for course characteristics (Items 33 to 35), 1 = *Much Less than Most Courses*, 2 = *Less than Most Courses*, 3 = *About Average*, 4 = *More than Most Courses*, 5 = *Much More than Most Courses*. For student characteristics, overall summary measures, and additional method items (Items 36-47), 1 = *Definitely False*, 2 = *More False than True*, 3 = *In Between*, 4 = *More True than False*, and 5 = *Definitely True*.

The statistics in Table 2 represent the same 47 items collected in different data collection formats and instruments, depending upon the year. Ratings from the fall of 2014 through the spring of 2015 were a mixture of IDEA paper and online administrations, using the 47-item original DF instrument, and were collected from courses with at least 10 student responses. Pilot data were collected exclusively on the Campus Labs platform, which was delivered online. In addition to the 47 original items, 13 pilot items were included in the DF instrument for students, and 6 pilot items were added to the OSF for instructors.

In examining Table 2, two key findings emerged. First, means in the 2015 pilot sample were generally slightly lower but nonetheless similar—especially those collected online—to means in the 2014-2015 research database. Second, pilot classes with a minimum of 10 responses had mean values very similar to those with a minimum of 5 responses. Also, classes with 10 or more responses had standard deviations very close to those from the 2014-2015 research dataset, which were consistently lower than those from classes with 5 or more

responses. Because smaller variability contributes to increased reliability, the decision was, therefore, made to restrict the pilot research sample to classes with at least 10 responses ( $n = 3,484$ ).

Table 2

*Means and Standard Deviations for Student Ratings of Existing Items on IDEA Diagnostic Feedback*

Item	2014-2015 Research dataset				2015 Pilot sample			
	Paper ( $n = 58,742$ )		Online ( $n = 29,689$ )		Classes with 5 responses or more ( $n = 6,405$ )		Classes with 10 responses or more ( $n = 3,484$ )	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Teaching methods								
1. Displayed personal interest in students	4.53	0.42	4.42	0.49	4.39	0.51	4.39	0.47
2. Helped students answer own questions	4.35	0.46	4.27	0.52	4.26	0.54	4.25	0.49
3. Scheduled work helpfully	4.41	0.42	4.37	0.46	4.34	0.50	4.34	0.45
4. Demonstrated importance of subject	4.51	0.40	4.42	0.46	4.43	0.47	4.42	0.43
5. Formed teams, discussion groups	3.97	0.81	4.02	0.77	3.94	0.79	3.93	0.76
6. Made clear how topics fit	4.43	0.44	4.36	0.49	4.37	0.51	4.37	0.46
7. Explained criticisms	4.19	0.50	4.17	0.54	4.16	0.57	4.13	0.52
8. Stimulated intellectual effort	4.21	0.49	4.19	0.52	4.17	0.55	4.15	0.50
9. Encouraged use of multiple resources	4.14	0.55	4.16	0.55	4.12	0.59	4.09	0.54
10. Explained clearly	4.34	0.54	4.24	0.59	4.26	0.61	4.26	0.56
11. Related to real life	4.43	0.49	4.39	0.51	4.36	0.54	4.36	0.49
12. Tests covered important points	4.43	0.42	4.39	0.45	4.34	0.49	4.35	0.44
13. Introduced stimulating ideas	4.31	0.50	4.28	0.54	4.25	0.56	4.25	0.51
14. Involved students in hands on activities	4.05	0.69	4.12	0.65	4.06	0.69	4.01	0.65
15. Inspired students to set high goals	4.12	0.55	4.11	0.57	4.09	0.60	4.06	0.55
16. Asked students to share experiences	4.06	0.67	4.11	0.66	4.05	0.69	4.02	0.65
17. Provided timely feedback	4.32	0.53	4.27	0.57	4.23	0.61	4.23	0.55
18. Asked students to help each other	4.15	0.54	4.14	0.56	4.07	0.60	4.05	0.55
19. Assessments required creativity	4.19	0.55	4.22	0.54	4.18	0.58	4.15	0.53
20. Encouraged student/faculty contact	4.23	0.53	4.23	0.54	4.14	0.59	4.14	0.53
Learning objectives								
21. Factual knowledge	4.27	0.44	4.24	0.45	4.12	0.49	4.11	0.45

22. Principles and theories	4.21	0.45	4.21	0.46	4.11	0.50	4.09	0.45
23. Applications	4.24	0.46	4.22	0.47	4.14	0.52	4.12	0.47
24. Professional skills, viewpoints	4.20	0.47	4.18	0.48	4.11	0.53	4.09	0.47
25. Team skills	3.83	0.68	3.85	0.63	3.75	0.69	3.73	0.63
26. Creative capacities	3.74	0.70	3.81	0.62	3.72	0.68	3.66	0.63
27. Broad liberal education	3.82	0.63	3.87	0.58	3.84	0.62	3.81	0.56
28. Communication skills	3.79	0.69	3.85	0.63	3.80	0.67	3.75	0.63
29. Find, use resources	3.98	0.52	4.02	0.50	3.94	0.56	3.91	0.50
30. Values development	3.88	0.61	3.93	0.58	3.86	0.63	3.84	0.57
31. Critical analysis	4.00	0.56	4.04	0.54	3.97	0.59	3.96	0.53
32. Interest in learning	4.06	0.51	4.06	0.52	3.99	0.56	3.96	0.51
Course characteristics								
33. Amount of reading	3.29	0.72	3.37	0.66	3.26	0.66	3.26	0.63
34. Amount of other work	3.53	0.52	3.52	0.51	3.40	0.52	3.37	0.48
35. Difficulty of subject matter	3.50	0.56	3.50	0.55	3.41	0.55	3.40	0.53
Student characteristics								
36. Strong desire to take the course	3.81	0.64	3.80	0.64	3.74	0.66	3.68	0.62
37. Worked harder on this course than most	3.74	0.52	3.75	0.52	3.70	0.54	3.65	0.51
38. Wanted this instructor	3.69	0.65	3.60	0.66	3.66	0.68	3.64	0.63
39. Wanted course regardless of instructor	3.52	0.54	3.54	0.54	3.43	0.55	3.39	0.48
43. Usually work hard on academic work	3.91	0.32	4.00	0.30	3.93	0.35	3.89	0.29
Summary measures								
40. Increase positive attitude toward field	4.06	0.56	4.01	0.59	4.01	0.60	3.98	0.56
41. Excellent instructor	4.35	0.56	4.26	0.61	4.31	0.61	4.31	0.57
42. Excellent course	4.15	0.55	4.10	0.59	4.12	0.60	4.09	0.57
Additional method items								
44. Used variety of evaluation methods	4.04	0.50	4.11	0.52	4.08	0.54	4.05	0.51
45. Expected students to take responsibility	4.39	0.31	4.46	0.31	4.40	0.36	4.37	0.32
46. High achievement standards	4.27	0.38	4.32	0.39	4.29	0.42	4.26	0.37
47. Used educational technology	4.23	0.51	4.33	0.47	4.24	0.53	4.24	0.48

### Analyses of the Pilot Research Dataset

Table 3 presents means and standard deviations for student ratings on the 13 pilot items. Existing objectives *factual knowledge* (Item 21,  $M = 4.11$ ) and *principles and theories* (Item 22,

$M = 4.09$ ), both of which measure acquiring basic cognitive information, consistently received the two highest ratings among the existing 12 objectives. In the same manner the highest ratings among pilot objectives were found for the objective *understanding subject matter* (Item 54,  $M = 4.14$ ), which was designed to combine the two above-mentioned objectives. On average, students reported less than substantial progress on developing *information literacy* (Item 53,  $M = 3.91$ ), *civic engagement* (Item 51,  $M = 3.86$ ), *understanding of diverse perspectives* (Item 49,  $M = 3.74$ ), *ethical reasoning* (Item 50,  $M = 3.74$ ), and *quantitative literacy* (Item 52,  $M = 3.70$ ). The most frequently observed pilot teaching methods were *encouraged self-reflection* (Item 56,  $M = 4.19$ ) and *provided meaningful feedback* (Item 58,  $M = 4.11$ ).

Table 3

*Means and Standard Deviations for Student Ratings of Pilot Items on IDEA Diagnostic Feedback (n = 3,484)*

Pilot item	<i>M</i>	<i>SD</i>
Learning objectives		
49. Diverse perspectives	3.74	0.63
50. Ethical reasoning	3.74	0.61
51. Civic engagement	3.86	0.56
52. Quantitative literacy	3.70	0.58
53. Information literacy	3.91	0.51
54. Understanding subject matter	4.14	0.45
Teaching methods		
55. Helped interpret subject matter	3.96	0.62
56. Encouraged self-reflection	4.19	0.51
57. Created service opportunities	3.98	0.57
58. Provided meaningful feedback	4.11	0.57
Student/course characteristics		
59. Self-efficacy	3.95	0.38
60. Background preparation	3.82	0.45
61. Amount of coursework	3.48	0.46

*Comparison of Faculty Ratings of Learning Objectives in the Pilot Sample and 2014-2015 Research Dataset*

We examined the extent to which faculty ratings of the relevance of the 12 existing learning objectives were similar between the pilot sample and the 2014-2015 research dataset. On the OSF, instructors were asked to indicate the relevance of each of the 12 existing learning objectives for the course they taught, using a 3-point scale ( $M = \textit{Minor or No Importance}$ ,  $I = \textit{Important}$ , and  $E = \textit{Essential}$ , coded as 1, 2, and 3 respectively). Only courses where the instructor had identified at least one existing learning objective as either important or essential were included in the analysis.

Table 4 presents the percentages of courses where faculty identified each of the existing learning objectives as *Minor or no importance*, *Important*, or *Essential* in the 2014-2015 research dataset. *Factual knowledge*, *applications*, and *principles and theories* were the most widely emphasized objectives, with about three quarters of courses identifying them as relevant (i.e., *Important* or *Essential*).

Table 4

*Descriptive Statistics for Faculty Ratings of the Relevance of 12 Existing Learning Objectives in the 2014-2015 Research Dataset*

Item	% Minor or no importance	% Important	% Essential	% Essential or Important	<i>N</i>
1. Factual knowledge	22.7	27.2	50.2	77.3	84,654
2. Principles and theories	27.1	31.6	41.3	72.9	84,124
3. Applications	23.8	34.6	41.6	76.2	84,543
4. Professional skills, viewpoints	47.1	27.6	25.3	52.9	82,376
5. Team skills	70.2	21.4	8.4	29.8	80,330
6. Creative capacities	81.9	10.5	7.5	18.1	79,663
7. Broad liberal education	76.5	13.6	9.9	23.5	80,083
8. Communication skills	56.9	23.6	19.5	43.1	81,451
9. Find, use resources	61.9	26.0	12.1	38.1	80,243
10. Values development	77.1	15.4	7.5	22.9	79,740
11. Critical analysis	51.8	25.2	23.0	48.2	81,475
12. Interest in learning	64.4	25.2	10.4	35.6	79,668

*Note.* Totals of percentages are not 100 for every objective because of rounding.

Table 5 presents descriptive statistics for each of the existing and proposed learning objectives in the pilot sample. Among the existing objectives, the rank of each based on the percentage of courses where it was identified as relevant nearly mirrors that in the 2014-2015 research dataset. *Applications*, *factual knowledge*, and *principles and theories* were the three most widely emphasized objectives, being relevant in about two thirds of courses. In general, the percentage of courses in which faculty identified each existing objective as relevant (i.e., *Essential* or *Important*) was lower in the pilot sample than in the 2014-2015 research dataset. Differences ranged from 6.7% to 15.3%. However, as in the research dataset, percentages in the pilot sample varied across objectives, indicating instructors were discriminating in their selections.

Table 5

*Descriptive Statistics for Faculty Ratings of the Importance of Existing and Proposed Learning Objectives in the Pilot Sample*

Item	% Minor or no importance	% Important	% Essential	% Essential or Important	N
Existing learning objectives					
1. Factual knowledge	33.9	25.6	40.5	66.1	3,484
2. Principles and theories	37.7	29.9	32.4	62.3	3,484
3. Applications	33.7	31.0	35.3	66.3	3,484
4. Professional skills, viewpoints	59.4	21.1	19.5	40.6	3,484
5. Team skills	80.3	14.6	5.2	19.7	3,484
6. Creative capacities	89.6	5.9	4.6	10.4	3,484
7. Broad liberal education	84.2	8.0	7.8	15.8	3,484
8. Communication skills	65.8	16.8	17.3	34.2	3,484
9. Find, use resources	75.2	15.7	9.1	24.8	3,484
10. Values development	83.8	10.7	5.6	16.2	3,484
11. Critical analysis	62.2	18.4	19.3	37.8	3,484
12. Interest in learning	79.7	14.4	5.8	20.3	3,484
Proposed learning objectives					
13. Diverse perspectives	78.5	12.9	8.6	21.5	3,199
14. Ethical reasoning	84.6	9.6	5.8	15.5	3,199
15. Civic engagement	82.4	11.8	5.9	17.7	3,199
16. Quantitative literacy	88.6	6.3	5.2	11.4	3,199
17. Information literacy	80.8	12.6	6.5	19.2	3,199
18. Understanding subject matter	58.5	20.1	21.4	41.5	3,199

*Note.* Totals of percentages are not 100 for every objective because of rounding.

Among the proposed objectives, *understanding subject matter* (Item 18 on OSF) was identified as relevant most widely, with more than two in five courses (41.5%) emphasizing it. *Quantitative literacy* (Item 16) was the least emphasized objective, identified as relevant in slightly over 10% of courses in the pilot sample.

### Selecting Objectives for the Revised Instruments

Although much had been accomplished in proposing which learning objectives to include in a revised instrument (see Benton et al., 2015), we aimed to confirm whether our plans would be validated in the pilot study. To begin, we review briefly the reasons why we included each of the objectives in the pilot study and the processes involved in arriving at those decisions.

As described in Benton et al. (2015), the process for revising the SRI began with focus groups of experienced IDEA users conducted at two professional conferences. We also surveyed IDEA staff to obtain their suggestions about possible changes. We next formed an updating team comprised of staff members from relevant units at the IDEA Center. The purpose of the team was to act as the decision-making body for all proposed changes. Two expert panels were recruited to review drafts of revisions and to provide feedback. Drafts of proposed items went through multiple reviews. After several rounds, we then contacted IDEA users and non-users and faculty to obtain additional comments on the suggested revisions. Finally, we conducted cognitive interviews with college students to test our proposed new items.

Another important step in the process was to review documents from higher education organizations involved in accreditation and outcomes assessment. We examined multiple sources, but took special care to align our changes with the AAC&U VALUE Rubrics and Lumina's Degree Qualifications Profile (DQP). Appendix A contains documents that show how each proposed learning objective aligns with outcomes specified by those organizations.

#### *Correlations Between Student and Faculty Ratings of Existing and Proposed Learning Objectives*

An indirect test of the validity of the IDEA SRI are the correlations between students' average ratings of progress on each objective at the course level and instructors' ratings of objective relevance. It is expected that students report more progress on objectives emphasized by their instructor than on other objectives. The highest correlations should be found in ratings of the same objectives if the following assumptions are true (Hoyt, 1973b, p. 376):

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

For the pilot study, we examined the extent to which correlations between faculty ratings of relevance and student ratings of progress followed expected patterns. Table 6 presents Pearson  $r$  correlations between faculty ratings of importance and student mean ratings of progress on existing and proposed objectives. The magnitudes and directions of the correlations are similar to those in Hoyt and Lee (2002). Fifteen of the 18 learning objectives exhibited the strongest positive correlations between faculty ratings of relevance and their corresponding student ratings of progress. For those 15 objectives, the correlations between instructor and student ratings of the same objective (indicated in bold font) were all statistically significant and ranged from .05 to .32 ( $M = 0.19$ ,  $SD = 0.07$ ). The mean coefficient for off-diagonal (i.e., noncorresponding) correlations was close to zero ( $M = 0.01$ ,  $SD = 0.08$ ). The strongest correlations were found for existing objectives on *team skills* ( $r = .26$ ,  $p < .001$ ), *creative capacities* ( $r = .26$ ,  $p < .001$ ), and *communication skills* ( $r = .32$ ,  $p < .001$ ), which is consistent with previous findings (Benton et al., 2015). These correlations provide evidence for criterion-related validity in that students tend to report greater progress on objectives stressed by their instructor.

Correlations between faculty and student ratings were noticeably low for two existing learning objectives: *applications* (FR3) and *interest in learning* (FR12). The correlation for FR12 ( $r = .05$ ,  $p < .01$ ) was consistent with what had been reported previously (Benton et al., 2015), which is why it was dropped. The one for *applications* was lower ( $r = .05$ ,  $p < .01$ ) than expected. Students were more likely to report more progress on the proposed objective *quantitative literacy* when their instructor emphasized *applications* ( $r = .18$ ,  $p < .001$ ). Given the limited sample size of the pilot study, we decided to withhold judgment and retain FR3 until additional data are available.

Table 6

*Correlations for Faculty Ratings of Relevance and Student Ratings of Progress on 12 Existing and 6 Proposed Learning Objectives*

Item	FR1	FR2	FR3	FR4	FR5	FR6	FR7	FR8	FR9	FR10	FR11	FR12	FR13	FR14	FR15	FR16	FR17	FR18
SR21	<b>.09***</b>	.03	.01	<b>.07***</b>	-.01	-.00	-.04*	-.04*	-.01	-.06***	-.04*	-.00	-.03	-.01	-.01	-.02	-.03*	-.02
SR22	<b>.06***</b>	<b>.05**</b>	.03	<b>.07***</b>	.01	-.01	-.05**	-.04**	-.01	-.03	-.02	.01	-.03	.01	-.00	-.01	-.03	-.03
SR23	-.06***	-.05**	<b>.05**</b>	<b>.13***</b>	<b>.04*</b>	.03	-.08***	.01	.01	.00	.00	.02	-.04*	<b>.03*</b>	<b>.04*</b>	-.02	.00	-.08***
SR24	-.04*	-.06**	<b>.04*</b>	<b>.16***</b>	<b>.05**</b>	<b>.07***</b>	-.06***	.02	.03	-.04**	-.03	.00	-.05**	.02	.03	-.02	.00	-.09***
SR25	-.14***	-.14***	<b>.09***</b>	<b>.15***</b>	<b>.26***</b>	<b>.07***</b>	-.04*	<b>.05**</b>	<b>.07***</b>	<b>.03*</b>	-.03	.03	.02	<b>.06***</b>	<b>.09***</b>	.03	<b>.04*</b>	-.11***
SR26	-.27***	-.26***	-.06***	<b>.11***</b>	<b>.10***</b>	<b>.26***</b>	<b>.15***</b>	<b>.25***</b>	<b>.08***</b>	.03	<b>.12***</b>	<b>.04*</b>	<b>.06***</b>	<b>.05**</b>	<b>.05**</b>	-.08***	<b>.11***</b>	-.13***
SR27	-.14***	-.17***	-.11***	.02	<b>.04*</b>	<b>.16***</b>	<b>.19***</b>	<b>.17***</b>	.03	<b>.03*</b>	<b>.11***</b>	<b>.06***</b>	<b>.15***</b>	.02	<b>.04*</b>	-.10***	.03	-.08***
SR28	-.26***	-.24***	-.07***	<b>.06***</b>	<b>.08***</b>	<b>.10***</b>	<b>.07***</b>	<b>.32***</b>	<b>.10***</b>	<b>.05**</b>	<b>.19***</b>	<b>.05**</b>	<b>.11***</b>	<b>.10***</b>	<b>.07***</b>	-.07***	<b>.12***</b>	-.13***
SR29	-.10***	-.11***	<b>.05**</b>	<b>.13***</b>	<b>.07***</b>	.02	-.07***	<b>.13***</b>	<b>.14***</b>	-.04*	<b>.07***</b>	.03	-.01	<b>.04*</b>	<b>.04*</b>	.01	<b>.11***</b>	-.13***
SR30	-.17***	-.13***	.01	<b>.09***</b>	<b>.08***</b>	<b>.06***</b>	-.00	<b>.08***</b>	<b>.06***</b>	<b>.16***</b>	<b>.11***</b>	<b>.08***</b>	<b>.09***</b>	<b>.14***</b>	<b>.13***</b>	-.09***	<b>.04*</b>	-.08***
SR31	-.17***	-.13***	-.00	<b>.05**</b>	<b>.05**</b>	.03	.02	<b>.15***</b>	<b>.07***</b>	<b>.06***</b>	<b>.21***</b>	<b>.07***</b>	<b>.06**</b>	<b>.11***</b>	<b>.07***</b>	-.05**	<b>.08***</b>	-.09***
SR32	-.09***	-.09***	.01	<b>.10***</b>	<b>.05**</b>	<b>.05**</b>	-.02	<b>.07***</b>	<b>.05**</b>	.03	<b>.07***</b>	<b>.05**</b>	.02	<b>.06**</b>	<b>.06**</b>	-.04*	<b>.04*</b>	-.08***
SR49	-.15***	-.14***	-.10***	.00	<b>.06**</b>	.03	<b>.12***</b>	<b>.17***</b>	<b>.04*</b>	<b>.11***</b>	<b>.19***</b>	<b>.08***</b>	<b>.28***</b>	<b>.11***</b>	<b>.13***</b>	-.13***	<b>.04*</b>	-.07***
SR50	-.17***	-.12***	.00	<b>.08***</b>	<b>.09***</b>	-.01	.00	<b>.11***</b>	<b>.07***</b>	<b>.15***</b>	<b>.20***</b>	<b>.08***</b>	<b>.17***</b>	<b>.20***</b>	<b>.16***</b>	-.08***	<b>.06**</b>	-.08***
SR51	-.12***	-.09***	<b>.05**</b>	<b>.14***</b>	<b>.09***</b>	.01	-.06***	<b>.04*</b>	<b>.07***</b>	<b>.11***</b>	<b>.09***</b>	<b>.06***</b>	<b>.09***</b>	<b>.14***</b>	<b>.16***</b>	-.05**	<b>.05**</b>	-.06***
SR52	<b>.06***</b>	<b>.05**</b>	<b>.18***</b>	<b>.15***</b>	<b>.07***</b>	-.08***	-.18***	-.07***	<b>.10***</b>	-.10***	-.03*	.01	-.10***	.01	.02	<b>.20***</b>	.03	-.08***
SR53	-.09***	-.10***	<b>.04*</b>	<b>.10***</b>	<b>.06***</b>	.00	-.04**	<b>.13***</b>	<b>.13***</b>	-.04*	<b>.12***</b>	<b>.04*</b>	.02	<b>.05**</b>	<b>.05**</b>	.03	<b>.13***</b>	-.11***
SR54	.02	-.01	.01	<b>.07***</b>	.01	.01	-.03	.00	.01	-.03	.02	.01	.02	.03	.03	-.01	.01	-.04*

Note. SR = student ratings of progress on Diagnostic Feedback instrument. FR = faculty ratings of relevance on Objective Selection Form. Boldface indicates correlations between student ratings of progress on and faculty ratings of the relevance of the same learning objective.  $N = 3,484$ .

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

The most surprising finding concerned the pilot objective *understanding subject matter* (FR18). The correlations between faculty ratings of this objective and student ratings of all 18 learning objectives were negative. The one between faculty and student ratings of FR18 was the weakest ( $r = -.04, p < .05$ ). FR18 was created to combine and replace the following two existing objectives:

1. FR1: *Gaining factual knowledge (terminology, classifications, methods, trends)*
2. FR2: *Learning fundamental principles, generalizations, or theories*

We hypothesize that the negative correlation for FR18 might have occurred because some faculty who emphasized gaining basic cognitive information in their classes selected FR1 and FR2 but not FR18 due to their similarity in wording. They might have considered the proposed objective redundant to the two existing objectives and ignored it. On the Campus Labs platform, missing values on OSF are automatically assigned a value of 1 for “minor or no importance.” Instructors who emphasized basic cognitive background in their classes, by virtue of selecting objectives FR1 and FR2, had students who probably reported much progress on *understanding subject matter*, even though their instructor did not select FR18. Therefore, the correlation between faculty ratings of relevance and student ratings of progress on FR18 was negligible ( $r = -.04, p < .05$ ). To test this hypothesis, we cross-tabulated faculty ratings of relevance on the FR1, FR2, and FR18. Table 7 presents the results of this analysis, summarized below.

Table 7

*Frequency of Instructor Ratings of Relevance on Learning Objectives Factual Knowledge, Principles and Theories, and Understanding Subject Matter in the Pilot Sample*

Ratings of relevance on <i>Understanding subject matter</i> (FR18)	Ratings of relevance on <i>Factual knowledge</i> (FR1)								
	Minor			Important			Essential		
	Ratings on <i>Principles &amp; theories</i> (FR2)			Ratings on <i>Principles &amp; theories</i> (FR2)			Ratings on <i>Principles &amp; theories</i> (FR2)		
	Minor	Important	Essential	Minor	Important	Essential	Minor	Important	Essential
Minor	512	156	77	153	199	103	148	167	357
Important	104	66	25	77	108	57	39	103	63
Essential	33	20	39	39	44	57	75	112	266
Total	649	242	141	269	351	217	262	382	686

Among faculty who rated FR1 and FR2 “Minor or of no importance” ( $n = 649$ ), 79% ( $n = 512$ ) rated FR18 “Minor or of no importance.” Of those who rated FR1 and FR2 “Important” ( $n = 351$ ), only 31% ( $n = 108$ ) rated FR18 “Important.” Finally, of faculty who rated FR1 and FR2 “Essential” ( $n = 686$ ), only 39% ( $n = 266$ ) rated FR18 “Essential.”

It was evident that approximately two-thirds of faculty who identified cognitive learning objectives as either important or essential—by virtue of their ratings on FR1 and FR2—assigned less importance to FR18, which supports our hypothesis for why there was a negligible correlation on that objective. Many instructors who emphasized cognitive learning objectives simply left FR18 blank.

### *Correlations Among Faculty Ratings of Relevance of Learning Objectives*

To further investigate which objectives to include in the DF we computed Pearson  $r$  correlations on faculty ratings of existing and proposed objectives (see Table 8). The general pattern of correlations among existing objectives was similar to that reported previously (Benton et al., 2015; Hoyt & Lee, 2002). The strongest correlation was between proposed objectives “Developing ethical reasoning and/or ethical decision making” (FR14) and “Learning to apply knowledge and skills to benefit others or serve the public good” (FR15),  $r = .45, p < .001$ . Instructors who intended to emphasize ethical reasoning also placed greater value on civic engagement, which intuitively makes sense. The second highest correlation was between existing objectives FR1 and FR2,  $r = .43, p < .001$ ; however, correlations between those two objectives and the proposed objective *understanding subject matter* (FR18) were not as strong ( $r = .25$  and  $.20, p < .001$ , respectively), confirming what was reported previously in Table 7.

Other correlations between existing and proposed objectives intuitively make sense. The proposed objective “Developing knowledge and understanding of diverse perspectives, global awareness, or other cultures” (FR13) was positively correlated with existing objective *broad liberal education* (FR7),  $r = .22, p < .001$ . And, the proposed objective on “Developing ethical reasoning and/or ethical decision making” (FR14) was most strongly correlated with the existing objective on *values development* (FR10),  $r = .36, p < .001$ , the objective it was intended to replace. Finally, the proposed objective “Learning how to find, evaluate, and use resources to explore a topic in depth” (FR17) was most strongly correlated with existing objective *find, use resources* (FR9),  $r = .36, p < .001$ , the objective it was intended to replace.

Table 8

*Correlations for Faculty Ratings of Relevance on 12 Existing and 6 Proposed Learning Objectives*

Item	FR1	FR2	FR3	FR4	FR5	FR6	FR7	FR8	FR9	FR10	FR11	FR12	FR13	FR14	FR15	FR16	FR17	FR18
FR1	—																	
FR2	.43***	—																
FR3	.05**	.14***	—															
FR4	-.02	-.03	.19***	—														
FR5	-.11***	-.09***	.12***	.19***	—													
FR6	-.15***	-.12***	-.07***	.09***	.16***	—												
FR7	-.10***	-.18***	-.18***	-.12***	.01	.29***	—											
FR8	-.26***	-.23***	-.14***	-.10***	.11***	.16***	.19***	—										
FR9	-.09***	-.09***	.16***	.15***	.17***	.03	-.01	.19***	—									
FR10	-.12***	-.04*	.07***	-.01	.13***	.04*	.08***	.07***	.10***	—								
FR11	-.22***	-.13***	-.01	-.13***	.03	-.01	.13***	.35***	.18***	.21***	—							
FR12	-.01	.01	.12***	.04*	.16***	.08***	.11***	.10***	.23***	.21***	.23***	—						
FR13	-.04*	-.04*	-.07***	-.01	.08***	.06**	.22***	.22***	.09***	.20***	.22***	.20***	—					
FR14	-.07***	.03	.12***	.14***	.18***	.02	.01	.10***	.17***	.36***	.24***	.25***	.30***	—				
FR15	-.05**	.00	.16***	.15***	.18***	.05**	.02	.05**	.18***	.26***	.15***	.22***	.35***	.45***	—			
FR16	.10***	.06**	.14***	.11***	.12***	-.01	-.03	-.02	.16***	.00	.03	.12***	.03	.12***	.12***	—		
FR17	-.08***	-.06***	.06***	.11***	.18***	.08***	.07***	.26***	.36***	.09***	.25***	.24***	.21***	.24***	.24***	.21***	—	
FR18	.25***	.20***	.02	.03	.02	-.02	.04*	-.08***	.04*	.04*	-.01	.14***	.17***	.15***	.17***	.19***	.20***	—

*Note.* FR = faculty ratings of relevance on Objective Selection Form.

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

*Factor Structure of Faculty Ratings of Relevance on Learning Objectives*

The next step in the process of selecting learning objectives was to investigate the underlying structure of faculty ratings of objective relevance. We conducted principal components analysis with Varimax rotation on faculty ratings of existing and proposed learning objectives on the OSF. Table 9 presents results of the analysis. Evidence for five factors emerged even though existing objectives *interest in learning* (FR12) and *critical analysis* (FR11) loaded about equally on the first and second factors. The first factor was comprised of proposed objectives *ethical reasoning* (FR14), *civic engagement* (FR15), and *diverse perspectives* (FR13), as well as existing objective *values development* (FR10). Thus, *ethical reasoning* shared a common latent trait with *values development*, which supported our plan to drop the latter (see Benton et al., 2015). Likewise, the proposed objective *information literacy* (FR7) loaded on a second factor with existing objective *find, use resources* (FR9), the item it was intended to replace. In similar fashion, existing objectives *factual knowledge* (FR1), *principles and theories* (FR2), and proposed objective *understanding subject matter* (FR18) comprised a third factor, which again gave validity to our plans to replace the former two with the latter.

Table 9

*Factor Loadings from Principal Components Factor Analysis with Varimax Rotation: Eigenvalues and Percentages of Variance for Faculty Ratings of Importance on 12 Existing and 6 Proposed Learning Objectives*

Faculty rating	Factor loading				
	1	2	3	4	5
FR14. Ethical reasoning	.74	.15	.02	-.07	.12
FR10. Values development	.69	-.08	-.14	-.04	-.01
FR15. Civic engagement	.68	.13	.07	-.01	.22
FR13. Diverse perspectives	.57	.15	.09	.34	-.14
FR12. Interest in learning	.40	.38	.09	.08	.04
FR17. Information literacy	.16	.73	.02	.10	.06
FR9. Find, use resources	.08	.66	-.14	-.17	.16
FR16. Quantitative literacy	-.05	.49	.31	-.06	.23
FR8. Communication skills	.08	.46	-.41	.30	-.25
FR11. Critical analysis	.36	.43	-.31	.00	-.42
FR1. Factual knowledge	-.11	-.06	.75	-.09	-.07
FR2. Principles, theories	.05	-.10	.66	-.23	-.08
FR18. Understanding subject matter	.17	.25	.64	.19	.00
FR7. Broad liberal education	.11	.04	-.06	.73	-.08
FR6. Creative capacities	-.01	.01	-.14	.64	.41
FR4. Professional skills, viewpoints	.05	.07	-.01	-.12	.71
FR5. Team skills	.19	.21	-.15	.10	.54
FR3. Applications	.15	.14	.07	-.50	.37
Eigenvalue	3.10	2.07	1.52	1.29	1.21

% of variance	17.23	11.47	8.44	7.16	6.74
---------------	-------	-------	------	------	------

Note.  $N = 3,484$ .

We proceeded to factor analyze the seven items retained from the existing instrument along with the six new proposed objectives. The initial solution produced five factors. However, *information literacy* (FR17) loaded on two factors, which created some ambiguity. We, therefore, forced a four-factor solution, the results of which appear in Table 10. The first factor addresses Developing General Life Skills and includes such objectives as *critical analysis*, *ethical reasoning*, and *diverse perspectives*. The second factor, Professional Skills, speaks to *skills and viewpoints needed by professionals* in the field, *applications* of course content, and *team skills*. Cultural/Creative Development pertains to development of a *broad, liberal education* and *creative capacities*. Finally, Learning Course-Specific Skills refers to *understanding subject matter* and *quantitative literacy*.

Table 10

*Factor Loadings from Principal Components Factor Analysis with Varimax Rotation: Eigenvalues and Percentages of Variance for Faculty Ratings of Importance on 13 Learning Objectives in Diagnostic Feedback 2016*

Item	Factor loading			
	General life skills	Professional skills	Cultural/creative development	Course-specific skills
FR11. Critical analysis	<b>.73</b>	-.17	-.03	-.21
FR8. Communication skills	<b>.60</b>	-.03	.30	-.37
FR14. Ethical reasoning	<b>.58</b>	.24	-.14	.30
FR13. Diverse perspectives	<b>.56</b>	-.10	.24	.32
FR17. Information literacy	<b>.54</b>	.21	.08	.22
FR15. Civic engagement	<b>.50</b>	.27	-.09	.39
FR4. Professional skills, viewpoints	-.09	<b>.68</b>	-.08	.14
FR5. Team skills	.20	<b>.65</b>	.13	-.04
FR3. Applications	.04	<b>.45</b>	-.48	.07
FR7. Broad liberal education	.15	-.12	<b>.73</b>	.07
FR6. Creative capacities	-.03	.40	<b>.71</b>	-.07
FR18. Understanding subject matter	.05	-.11	.09	<b>.80</b>
FR16. Quantitative literacy	.08	.23	-.08	<b>.45</b>
Eigenvalue	2.49	1.75	1.29	1.14
% of variance	19.18	13.47	9.89	8.80

Note.  $N = 3,484$ . Boldface indicates highest factor loadings.

#### *Factor Structure of Student Ratings of Progress on Learning Objectives*

To better understand the underlying structure of student ratings of progress, we performed principal components analysis with Varimax rotation on existing and proposed learning objectives. Table 11 presents factor loadings, eigenvalues, and percentages of variance explained. As with faculty ratings of relevance, student ratings of proposed objectives *ethical reasoning* (Item 50), *civic engagement* (Item 51), and *diverse perspectives* (Item 49) shared a common latent trait with existing

objective *values development* (Item 30), which boosted our confidence that the latter objective could be discarded. Similarly, existing objectives *factual knowledge* (Item 21) and *principles and theories* (Item 22) loaded on the same factor with proposed objective *understanding subject matter* (Item 54). So, again, the proposed objective shared common variance with the two objectives it was intended to replace. Finally, the proposed objective *information literacy* (Item 53) loaded on the same factor with existing objective *find, use resources* (Item 29), the one it was intended to replace.

Table 11

*Factor Loadings from Principal Components Factor Analysis with Varimax Rotation: Eigenvalues and Percentages of Variance for Student Ratings of Progress on 12 Existing and 6 Proposed Learning Objectives*

Item	Factor loading	
	1	2
28. Communication skills	.87	.33
49. Diverse perspectives	.86	.28
26. Creative capacities	.83	.31
50. Ethical reasoning	.82	.39
30. Values development	.80	.47
27. Broad liberal education	.77	.43
31. Critical analysis	.74	.56
51. Civic engagement	.73	.54
25. Team skills	.63	.45
21. Factual knowledge	.34	.88
22. Principles and theories	.37	.88
54. Understanding subject matter	.43	.83
24. Professional skills, viewpoints	.47	.82
23. Applications	.48	.82
52. Quantitative literacy	.25	.76
32. Interest in learning	.63	.71
29. Find, use resources	.60	.69
53. Information literacy	.60	.68
Eigenvalues	13.57	1.31
% of variance	75.41	7.30

Note.  $N = 3,484$ .

We then factor analyzed student ratings of progress on the 13 objectives retained in the *Diagnostic Feedback 2016* instrument (see Table 12). The first factor was similar to the faculty ratings dimension of Developing General Life Skills that are useful throughout the lifespan and not specific to a course. Such skills included *diverse perspectives* (Item 49), *communication skills* (Item 28), and *creative capacities* (Item 26). The second component involved Learning Course-Specific Skills, such as *quantitative literacy* (Item 52), *understanding subject matter* (Item 54), and *applications* (Item 23). Both subscales of learning objectives had high internal consistency as demonstrated by the Cronbach's alpha coefficients reported in Table 12. Institutions interested in assessing students'

progress on obtaining General Life Skills and Course-Specific Skills could create such subscales by computing student mean ratings of progress on objectives loading on each of those factors.

Table 12

*Factor Loadings from Principal Components Factor Analysis with Varimax Rotation: Eigenvalues and Percentages of Variance for Student Ratings of Progress on Learning Objectives on IDEA2*

Item	Factor loading	
	General life skills	Course-specific skills
49. Diverse perspectives	<b>.86</b>	.28
28. Communication skills	<b>.86</b>	.34
26. Creative capacities	<b>.84</b>	.32
27. Broad liberal education	<b>.82</b>	.38
50. Ethical reasoning	<b>.79</b>	.44
31. Critical analysis	<b>.73</b>	.57
51. Civic engagement	<b>.69</b>	.60
25. Team skills	<b>.58</b>	.52
52. Quantitative literacy	.17	<b>.87</b>
54. Understanding subject matter	.45	<b>.80</b>
23. Applications	.49	<b>.80</b>
24. Professional skills, viewpoints	.48	<b>.79</b>
53. Information literacy	.56	<b>.74</b>
Eigenvalues	9.62	1.00
% of variance	74.04	7.72
Cronbach's alpha	.96	.95

*Note.*  $N = 3,484$ . Boldface indicates highest factor loadings.

#### *Student Progress on Relevant Objectives*

IDEA believes the best measure of teaching effectiveness is average student progress on relevant objectives. Converted scores (i.e., T-scores) represent an instructor's mean student rating relative to the overall mean for classes identifying a particular objective as either important or essential. Table 13 presents means and standard deviations for student ratings on each of the existing and proposed objectives identified as relevant in the course. The statistics for proposed objectives will be used to create comparative scores for the overall pilot dataset in the initial rollout of class reports in Spring 2016. [The formula for creating T Scores is as follows:](#)

$$T \text{ Score} = 50 + \frac{\text{Obtained Mean} - \text{Grand Mean}}{SD} \times 10$$

where the Grand Mean is the National mean of the IDEA database and SD is the National standard deviation. We will continue updating the statistics annually as we collect more data in the future.

Deleted: .

Table 13

*Means and Standard Deviations for Student Ratings on Relevant Learning Objectives in the Pilot Sample*

Item	<i>M</i>	<i>SD</i>	<i>N</i>
Existing learning objectives			
21. Factual knowledge	4.14	0.44	2,302
22. Principles and theories	4.12	0.44	2,171
23. Applications	4.13	0.49	2,309
24. Professional skills, viewpoints	4.18	0.47	1,415
25. Team skills	4.07	0.50	687
26. Creative capacities	4.15	0.47	364
27. Broad liberal education	4.06	0.51	550
28. Communication skills	4.03	0.50	1,190
29. Find, use resources	4.03	0.48	864
30. Values development	4.06	0.48	565
31. Critical analysis	4.10	0.46	1,316
32. Interest in learning	4.02	0.51	706
Proposed learning objectives			
49. Diverse perspectives	4.07	0.47	689
50. Ethical reasoning	4.01	0.47	495
51. Civic engagement	4.06	0.49	565
52. Quantitative literacy	4.00	0.52	366
53. Information literacy	4.03	0.47	614
54. Understanding subject matter	4.13	0.44	1,328

*Note.* Only classes where the instructor identified the objective as either “Essential” or “Important” were included in each analysis.

We were especially interested in comparing means of proposed objectives with those they were intended to replace. For example, the mean for *understanding subject matter* (Item 54,  $M = 4.13$ ) is equal to the average of *factual knowledge* (Item 21,  $M = 4.14$ ) and *principles and theories* (Item 22,  $M = 4.12$ ), the objectives it was designed to replace. Likewise, the mean for *information literacy* ( $M = 4.03$ ) is equal to the average of *find, use resources* ( $M = 4.03$ ) and *interest in learning* ( $M = 4.02$ ), items it will replace. Finally, the means of proposed objectives *diverse perspectives* (Item 49,  $M = 4.07$ ), *ethical reasoning* (Item 50,  $M = 4.01$ ), and *civic engagement* (Item 51,  $M = 4.05$ ) are comparable to the mean of existing objective *values development* (Item 30,  $M = 4.06$ ), the objective they are intended to replace.

The preponderance of the evidence convinced us to include all six proposed objectives in the revised DF form. First, the proposed objectives conceptually align with essential general learning outcomes defined by several higher education organizations. Second, correlations between faculty ratings of relevance and student ratings of progress are generally highest for the same objectives. Third, faculty ratings of relevance and student ratings of progress on those objectives share underlying dimensions with those they will replace. Fourth, mean ratings of

student progress on proposed objectives are comparable to those of the objectives they will replace. Therefore, the following 13 learning objectives are included in both the revised DF and LE:

1. Gaining a basic understanding of the subject (e.g., factual knowledge, methods, principles, generalizations, theories)
2. Developing knowledge and understanding of diverse perspectives, global awareness, or other cultures
3. Learning to *apply* course material (to improve thinking, problem solving, and decisions)
4. Developing specific skills, competencies, and points of view needed by professionals in the field most closely related to this course
5. Acquiring skills in working with others as a member of a team
6. Developing creative capacities (inventing; designing; writing; performing in art, music, drama, etc.)
7. Gaining a broader understanding and appreciation of intellectual/cultural activity (music, science, literature, etc.)
8. Developing skill in expressing myself orally or in writing
9. Learning how to find, evaluate, and use resources to explore a topic in depth
10. Developing ethical reasoning and/or ethical decision making
11. Learning to *analyze* and *critically evaluate* ideas, arguments, and points of view
12. Learning to apply knowledge and skills to benefit others or serve the public good
13. Learning appropriate methods for collecting, analyzing, and interpreting numerical information

#### **Selecting Teaching Methods for IDEA2 *Diagnostic Feedback***

As was the case with deciding which learning objectives to include in the revised DF, we consulted multiple sources of information, as explained in Benton et al. (2015), when considering revisions to teaching methods. Focus groups made suggestions about which items to drop and which new items to add; the expert panel members assisted in writing and revising new items; and the IDEA Updating Team evaluated those inputs. In addition, we aligned existing and proposed teaching methods with engagement indicators from NSSE (see Appendix C). Based on those several information sources, the following four proposed methods were included in the pilot study on the DF:

55. Helped students to interpret subject matter from diverse perspectives (e.g., different cultures, religions, genders, political views)
56. Encouraged students to reflect on and evaluate what they have learned
57. Created opportunities for students to apply course content outside the classroom
58. Provided meaningful feedback on students' academic performance

#### *Relationships between Teaching Methods and Relevant Learning Objectives*

An assumption of IDEA SRI is that the relationships between teaching methods and

relevant learning objectives are distinctive for each objective. For example, “Made it clear how each topic fit into the course” (Item 6) and “Stimulated students to intellectual effort beyond that required by most courses” (Item 8) are strongly associated with student progress on cognitive learning objectives *factual knowledge* and *principles and theories*; however, they are less important for acquiring *team skills* and developing *creative capacities*. *Team skills* and *creative capacities* benefit more from “Involved students in hands on projects” (Item 14) and “Inspired students to set and achieve goals which really challenged them” (Item 15).

To investigate which teaching methods were 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 from a single model (Hoeting, Madigan, Raftery, & Volinsky, 1999). BMA can provide estimated probabilities that student ratings of the frequency of each teaching method are associated with progress on a given learning objective. The Schwartz Bayesian Criterion (SBC) was used for model selection among the finite set of models (2 to the  $k$ th power, where  $k$  is the number of explanatory variables). We selected the best 100 models, based on the SBC criterion. Separate analyses were conducted on each learning objective, including only classes where the instructor rated the learning objective as relevant to the course. Tables of estimated probabilities and regression parameters (weighted coefficients) are presented for each learning objective in Appendix D. Table 14 summarizes the significant explanatory variables (indicated by item number on the DF) included in the “best” full models for each objective in medium size classes with an enrollment of 15 to 34 students. We will expand the analyses to small, large, and very large classes as more data are collected.

Table 14

*Teaching Methods Significantly Related with Progress on Learning Objectives in Medium-sized Classes (15 - 34)*

Learning objective	Relevant teaching methods
23. Learning to apply course material (to improve thinking, problem solving, and decisions)	8 (4, 11, 12, 15, 56)
24. Developing specific skills, competencies, and points of view needed by professionals in the field most closely related to this course	4, 6 (8, 14, 15, 18, 57)
25. Acquiring skills in working with others as a member of a team	5, 15 (14, 55)
26. Developing creative capacities (writing, inventing, designing, performing in art, music, drama, etc.)	15, 19
27. Gaining a broader understanding and appreciation of intellectual/cultural activity (music, science, literature, etc.)	13, 15, 55
28. Developing skill in expressing myself orally or in writing	15, 19, 55, 58 (9)
31. Learning to analyze and critically evaluate ideas, arguments, and points of view	8, 56 (13, 19, 55)
49. Developing knowledge and understanding of diverse perspectives, global awareness, or other cultures	55
50. Developing ethical reasoning and/or ethical decision making	8, 11, 55
51. Learning to apply knowledge and skills to benefit others or serve the public good	11, 55, 56, 57
52. Learning appropriate methods for collecting, analyzing, and interpreting numerical information	20
53. Learning how to find, evaluate, and use resources to explore a topic in depth	9, 15, 57, 58
54. Gaining a basic understanding of the subject (e.g., factual knowledge, methods, principles, generalizations, theories)	6, 8, 13, 56, 57

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

## Teaching Methods

- |   |   |
|---|---|
| 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   | 18. Asked students to help each other understand ideas or concepts  |
| 6. Made it clear how each topic fit into the course   | 19. Gave projects, tests, or assignments that required original or creative thinking  |
| 8. Stimulated students to intellectual effort beyond that required by most courses  | 20. Encouraged student-faculty interaction outside of class (office visits, phone calls, email, etc.)                                     |
| 9. Encouraged students to use multiple resources (e.g., data banks, library holdings, outside experts) to improve understanding | 55. Helped students to interpret subject matter from diverse perspectives (e.g., different cultures, religions, genders, political views) |
| 11. Related course material to real life situations   | 56. Encouraged students to reflect on and evaluate what they have learned   |
| 12. Gave tests, projects, etc. that covered the most important points of the course   | 57. Created opportunities for students to apply course content outside the classroom  |
| 13. Introduced stimulating ideas about the subject  | 58. Provided meaningful feedback on students' academic performance  |

In deciding which teaching methods to retain in the revised instrument, we considered the following criteria:

1. How often was the teaching method a significant predictor in the models?
2. How many times was the teaching method among the top 3 predictors in a model?
3. Was it an important teaching method for aligning with NSSE engagement indicators?
4. Was it an important teaching method for forming an *a priori* subscale? Hoyt and Lee (2002) posited five subscales: Stimulating Student Interest, Fostering Student Collaboration, Establishing Rapport, Encouraging Student Involvement, and Structuring Classroom Experiences.

Using those criteria, we decided not to include the following three teaching methods in the revised DF:

1. “Displayed a personal interest in students and their learning.” This item will be retained in *IDEA Teaching Essentials* because of its relationship with the two overall summary measures.
3. “Scheduled course work (class activities, tests, projects) in ways which encouraged students to stay up-to-date in their work”
12. “Gave tests, projects, etc. that covered the most important points of the course”

#### *Factor Structure of Student Ratings of Teaching Methods*

In order to understand the underlying structure of student ratings of teaching methods, we conducted principal components analysis with Varimax rotation on 17 existing and 4 proposed items. Table 15 presents factor loadings, eigenvalues, and percentages of variance explained for all items loading on two factors. The first factor appears to represent Instructor-Centered Teaching and includes existing methods such as *explained material clearly* (Item 10), *demonstrated importance of subject* (Item 4), and *made it clear how topics fit* (Item 6). Proposed methods loading on this factor were *provided meaningful feedback* (Item 58) and *encouraged self-reflection* (Item 56). The second factor is more pertinent to Student-Centered Teaching and includes the existing methods of *formed teams* (Item 5), *involved students in hands-on activities* (Item 14), and *asked diverse students to share ideas* (Item 16). It is worth noting that the proposed methods pertaining to *created service opportunities* (Item 57) and *helped interpret subject matter from diverse perspectives* (Item 55) exhibited loadings on both factors.

Table 15

*Factor Loadings from Principal Components Factor Analysis with Varimax Rotation: Eigenvalues and Percentages of Variance for Student Ratings of 17 Existing and Four Proposed Teaching Methods*

Item	Factor loading	
	1	2
10. Explained clearly	<b>.89</b>	.29
4. Demonstrated importance of subject	<b>.86</b>	.35
6. Made clear how topics fit	<b>.86</b>	.35
2. Helped students answer own questions	<b>.85</b>	.41

13. Introduced stimulating ideas	<b>.83</b>	.44
58. Provided meaningful feedback	<b>.82</b>	.40
8. Stimulated intellectual effort	<b>.81</b>	.42
17. Provided timely feedback	<b>.80</b>	.24
7. Explained criticisms	<b>.79</b>	.46
20. Encouraged student/faculty contact	<b>.77</b>	.40
56. Encouraged self-reflection	<b>.76</b>	.52
15. Inspired students to set high goals	<b>.74</b>	.55
11. Related to real life	<b>.70</b>	.43
57. Created service opportunities	.64	.63
55. Helped interpret subject matter from diverse perspectives	.63	.57
5. Formed teams, discussion groups	.15	<b>.87</b>
14. Involved students in hands on activities	.32	<b>.83</b>
16. Asked students to share experiences	.45	<b>.78</b>
19. Assessments required creativity	.55	<b>.71</b>
18. Asked students to help each other	.56	<b>.70</b>
9. Encouraged use of multiple resources	.53	<b>.64</b>
Eigenvalue	15.60	1.31
% of variance	74.07	5.95

Note.  $N = 3,484$ . Boldface indicates highest factor loadings.

Item 58, “Provided meaningful feedback on students’ academic performance,” was created to replace two existing teaching methods: “Explained the reasons for criticisms of students’ academic performance” (Item 7), and “Provided timely and frequent feedback on tests, reports, projects, etc. to help students improve” (Item 17). Input from focus groups and the expert panels suggested these latter two items were outdated, because one focused on “criticism” rather than feedback and the other emphasized timeliness and frequency when the key issue should be meaningfulness. As indicated in Table 15 all three methods loaded on the first factor, which means they share an underlying latent trait. Moreover, Item 58 is highly correlated with Items 7 ( $r = .89, p < .001$ ) and 17 ( $r = .83, p < .001$ ). This provides empirical evidence to support our decision to replace Items 7 and 17 with the Item 58 in the revised DF. Consequently, student ratings on teaching methods 7 and 17 were not included in the analyses described in the following paragraph.

In order to define the underlying structure of the 19 teaching methods included in the revised DF we conducted principal component analysis with Varimax rotation. The initial eigenvalues revealed two dimensions. Following Varimax rotation, the first factor appeared to represent methods that are Instructor Centered. As shown in Table 16, it was comprised of such methods as *explained clearly* (Item 10), *demonstrated significance* (Item 4), *made clear how topics fit* (Item 6), *provided meaningful feedback* (Item 58), and *encouraged out-of-class contact* (Item 20). The second factor, which pertained to Student-Centered methods, included such behaviors as *formed teams* (Item 5), *involved students in hands-on projects* (Item 14), and *asked diverse students to share ideas* (Item 16). As shown in Table 16, both factors had high reliability as subscales ( $\alpha = .98$  and  $.94$  respectively). We decided, therefore, to include all other existing

methods (except Items 7 and 17) and proposed teaching methods in the revised DF.

Table 16

*Factor Loadings from a Principal Components Factor Analysis with Varimax Rotation: Eigenvalues and Percentages of Variance for Student Ratings of Teaching Methods on IDEA2 Diagnostic Feedback*

Item	Factor loading	
	Instructor-centered	Student-centered
10. Explained clearly	<b>.90</b>	.27
4. Demonstrated significance	<b>.88</b>	.32
6. Made clear how topic fits	<b>.88</b>	.33
13. Introduced stimulating ideas	<b>.86</b>	.41
2. Helped students answer own questions	<b>.85</b>	.39
8. Stimulated intellectual effort	<b>.82</b>	.41
58. Provided meaningful feedback	<b>.79</b>	.41
56. Encouraged self-reflection	<b>.78</b>	.50
20. Encouraged out-of-class contact	<b>.76</b>	.40
15. Inspired ambitious goals	<b>.75</b>	.54
11. Related to real life	<b>.74</b>	.39
57. Created service opportunities	<b>.66</b>	.61
55. Helped interpret subject matter	<b>.65</b>	.54
5. Formed teams	.15	<b>.88</b>
14. Involved in hands on	.34	<b>.82</b>
16. Asked diverse students to share ideas	.48	<b>.76</b>
18. Asked students to help others	.56	<b>.71</b>
19. Required originality	.56	<b>.71</b>
9. Encouraged using multiple resources	.54	<b>.64</b>
Eigenvalues	14.19	1.24
% of variance	74.66	6.54
Cronbach's alpha	.98	.94

Note.  $N = 3,484$ . Boldface indicates highest factor loadings.

### Overall Summary Measures

IDEA SRI uses two summary measures to assess students' overall impressions of the instructor and the course: "Overall, I rate this instructor an excellent teacher" (Item 41) and "Overall, I rate this course as excellent" (Item 42). Students respond to these items, using the scale of 1 = *Definitely False*, 2 = *More False than True*, 3 = *In Between*, 4 = *More True than False*, and 5 = *Definitely True*. As shown in Table 2 the means and standard deviations for *excellent instructor* ( $M = 4.31$ ,  $SD = 0.57$ ) in the pilot were almost identical to those of the 2014-2015 research dataset ( $M = 4.32$ ,  $SD = 0.58$ ). Similarly, descriptive statistics for *excellent course* in the pilot sample ( $M = 4.09$ ,  $SD = 0.57$ ) were highly similar to those in the 2014-2015

research dataset ( $M = 4.13$ ,  $SD = 0.57$ ). Both summary measures will be retained in their current form. Item 40, “As a result of taking this course, I have more positive feelings toward this field of study,” will be removed because it provides ambiguous results (Benton et al., 2015).

### Selecting Variables for the Adjusted Score Formulas

Not all courses are alike. Class size varies. Students vary in their background preparations, motivation to take the course, and general work habits. Moreover, subject matter difficulty depends upon its content, structure, and student background preparation. One of the hallmarks of IDEA SRI is that ratings are adjusted for student and course characteristics that are beyond the instructor’s control. Mean class scores on several items are used to compute adjusted scores based on regression models produced for each learning objective and the two overall summary measures. The adjustments are intended to level the playing field to enable fairer comparisons among instructors.

#### Course Characteristics

Two course characteristics contribute to the existing adjusted score models: *workload* and *subject matter difficulty*. Instructors have some control over these factors in the amount of work and reading they assign and how much they stimulate student intellectual effort. In the existing version of the DF, two items are used to measure workload—“Amount of reading” (Item 33) and “Amount of work in other (non-reading) assignments” (Item 34). These two items, along with a teaching method, “Stimulated students to intellectual effort beyond that required by most courses” (Item 8), were used to create a residual score for “Difficulty of subject matter” (Item 35) (Hoyt & Lee, 2002). For the pilot instrument, we added a new item, “Amount of coursework” (Item 61), which was intended to replace the existing workload items. For all three items students responded using 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*. In the pilot sample, the mean rating for Item 61 was correlated moderately with Item 33 ( $r = .55$ ,  $p < .001$ ), and highly with Item 34 ( $r = .86$ ,  $p < .001$ ).

To investigate whether Item 61 could suffice as a measure of workload, we regressed “Difficulty of subject matter” (Item 35) on Item 61 on the first step, and then entered both Items 33 and 34 on the second step. The results are presented in Table 17. After controlling for Item 61, Items 33 and 34 contributed an additional less than 1% of explained variance. For this reason, the decision was made to include Item 61 in the revised instrument but to not retain the other two items.

Table 17

*Hierarchical Regression Analysis Summary for Workload Variables Predicting Difficulty of Subject Matter (N = 3,484)*

Step and predictor variable	<i>B</i>	<i>SE B</i>	$\beta$	$R^2$	$\Delta R^2$
Step 1:				.47***	
Amount of coursework	0.77	0.01	.68***		
Step 2:				.47***	.00
Amount of coursework	0.71	0.03	.63***		

Amount of reading	0.08	0.01	.09***
Amount of non-reading assignments	0.01	0.03	.01

\*\*\*  $p < .001$

#### *Student Characteristics*

The following five items measure student attitudes and behaviors in the course, using a 5-point Likert scale (1 = *Definitely False* and 5 = *Definitely True*):

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.
43. As a rule, I put forth more effort than other students on academic work.

Mean responses to *course motivation* (Item 39) and *work habits* (Item 43) are the two extraneous variables that historically carry the greatest weight in the regression models for computing adjusted scores (Hoyt & Lee, 2002). Mean scores on *effort* (Item 37) contribute to a *residual effort* score ( $E_N$ ) after controlling for *stimulating intellectual effort* (Item 8), *amount of reading* (Item 33), and *amount of non-reading assignments* (Item 34). Items 36 (“I had a strong desire to take this course”) and 38 (“I really wanted to take a course from this instructor”) were included to create a residual score on *other motivation*; however, the residual added so little variance to the regression models that Hoyt and Lee (2002) did not include it in any adjusted score formulas. Therefore, Items 36 and 38 were not included in the current regression modeling for adjusted scores, and they will not be included in a revised DF.

#### *Residual Difficulty and Effort Scores*

Hoyt and Lee (2002) created two residual scores,  $D_N$  and  $E_N$  respectively, for the adjusted score formulas.  $D_N$  and  $E_N$  represent student perception of course difficulty and effort after removing an instructor’s influence. For the current analysis, we measured the instructor’s influence on course difficulty and student effort, using mean student responses to the “Amount of coursework” (Item 61) and the teaching method *stimulating intellectual effort* (Item 8). By regressing ratings of “Difficulty of subject matter” (Item 35) on mean responses to Items 8 and 61 we found the following coefficients and constant:

$$X_{35} = .004 \times X_8 + .773 \times X_{61} + .691. \quad (1)$$

The residual for subject matter difficulty is then formed by:

$$D_N = \text{Mean of } X_{35} - \text{Predicted } X_{35}. \quad (2)$$

Next, we regressed *effort* (Item 37) on Item 8 and Item 61 to create the following formula:

$$\text{Predicted } X_{37} = .294 \times X_8 + .707 \times X_{61} + -.033. \quad (3)$$

The residual for effort was:

$$E_N = \text{Mean of } X_{37} - \text{Predicted } X_{37}. \quad (4)$$

The above formulas show that the more coursework an instructor required, and the more frequently students observed the instructor stimulated their intellectual effort, the more difficult they perceived the course to be and the more effort they put forth. The amount of variance explained in predicting subject matter difficulty ( $R^2 = .469$ ) and effort ( $R^2 = .605$ ) was close to that of Hoyt and Lee (2002),  $R^2 = .371$  and  $.635$ , respectively, but it was achieved with one less variable.

#### *Background Preparation and Self-Efficacy*

As reported in Benton et al. (2015), two new items were added to the pilot to examine the extent to which they contribute to the explained variance in the adjusted score formulas: “My background prepared me well for this course’s requirements” (*background preparation*, Item 60) and “When this course began I believed I could master its content” (*self-efficacy*, Item 59). *Background preparation* (BP) has been an item for years on *Learning Essentials* (previously named the *Short Form*). Benton et al. (2015) found that BP added meaningful and significant explained variance beyond *course motivation* and *work habits* to the adjusted score formulas. *Self-efficacy* (SE) was developed with the collaboration of expert panel members and is intended to measure students’ beliefs in their ability to succeed in the course.

Table 18 shows the Pearson  $r$  correlations between the independent variables in the adjusted score regression models.

Table 18

#### *Correlations for Adjusted Score Model Independent Variables*

Variable	CM	WH	Enroll	D <sub>N</sub>	E <sub>N</sub>	BP	SE
CM	—						
WH	.30***	—					
Enroll	-.09***	-.06***	—				
D <sub>N</sub>	.02	.06**	.14***	—			
E <sub>N</sub>	.33***	.29***	-.04	.44***	—		
BP	.41***	.39***	-.12***	-.25***	.05**	—	
SE	.43***	.35***	-.06**	-.31***	.02	.75***	—

*Note.* CM = Course Motivation, WH = Work Habits, D<sub>N</sub> = Difficulty unrelated to the instructor, E<sub>N</sub> = Effort unrelated to the instructor, BP = Background Preparation, SE = Self-Efficacy. Classes with response rates less than 50% were excluded.  $N = 2,997$ .

\*\*  $p < .01$ . \*\*\*  $p < .001$

Table 19 presents constants and standardized regression coefficients for the seven variables included in the regression models. Courses with response rates lower than 50% were excluded from the analyses. For each of 13 regression models average student progress on the respective learning objective was the dependent variable. In each model, only classes where the instructor identified the objective as important or essential were included. *Background*

*preparation* was the most potent predictor in the models except for the learning objectives on *broad liberal education* (Item 27), *diverse perspectives* (Item 49) and *civic engagement* (Item 51). Students' beliefs about their level of preparation (BP) had a strong positive effect on their perceived level of progress on most objectives and their overall ratings of the instructor and course. *Work habits* was the most important variable for making progress on *broad liberal education* (Item 27) and *diverse perspectives* (Item 49), and had the second highest regression coefficient in all other models except the proposed learning objective *understanding subject matter* (Item 54) and *excellence of the course*. *Course motivation* was relatively more important for ratings on *civic engagement* and *excellence of the course*, and it was an important component of several other models. So, *background preparation*, *work habits*, and *course motivation* stand out as three variables to definitely retain in the models. Classes comprised primarily of students who report good background preparation, sound work habits, and a strong desire to take the course regardless of who taught it tend to have higher ratings on course objectives and overall summary measures.

Table 19

*Regression Analyses Predicting Student Ratings on Relevant Learning Objectives and Summary Measures with Course Motivation, Work Habits, Enrollment, Difficulty Residual, Effort Residual, Background Preparation, and Self-efficacy*

Criterion	Constant	Standardized Regression Coefficients							R <sup>2</sup>	N
		CM	WH	Enroll	D <sub>N</sub>	E <sub>N</sub>	BP	SE		
23. Applications	0.801	.160	.212	—	—	-.063	.353	—	.308	1,801
24. Professional skills, viewpoints	0.740	.167	.249	—	—	—	.256	—	.314	1,135
25. Team skills	0.903	.098	.294	—	—	—	.353	-.128	.274	548
26. Creative capacities	1.998	.135	.172	-.164	—	.154	.195	—	.254	308
27. Broad liberal education	1.264	—	.247	-.141	—	-.122	.139	—	.195	432
28. Communication skills	1.674	.068	.235	-.150	-.107	--	.391	-.181	.243	1,007
31. Critical analysis	1.486	.092	.213	-.075	.089	-.074	.350	—	.231	1,067
49. Diverse perspectives	1.876	.132	.177	-.101	—	-.124	--	—	.137	554
50. Ethical reasoning	1.335	.175	.176	—	—	-.128	.290	—	.230	392
51. Civic engagement	0.380	.261	.212	—	—	-.179	.237	—	.313	463
52. Quantitative literacy	0.870	—	.295	—	—	—	.478	-.243	.287	287
53. Information literacy	1.317	—	.285	-.115	—	—	.286	—	.207	510
54. Understanding subject matter	1.129	.188	.177	—	—	-.068	.322	—	.297	1,101
41. Excellence of teacher	1.330	.053	.167	—	—	-.081	.311	—	.184	2,777
42. Excellence of course	0.398	.254	.139	—	—	—	.293	.051	.337	2,777

*Note.* CM = Course Motivation, WH = Work Habits, Enroll = enrollment, D<sub>N</sub> = Difficulty unrelated to the instructor, E<sub>N</sub> = Effort unrelated to the instructor, BP = Background Preparation, SE = Self-Efficacy. Classes with response rates less than 50% were excluded. Each model for learning objectives only included classes where the instructor rated the objective as relevant to the course.

Of relatively less importance were class size, residual scores on difficulty (D<sub>N</sub>) and effort (E<sub>N</sub>), and student self-efficacy. The weak influence of enrollment size and the residual scores is consistent with previous findings (Benton et al., 2015; Hoyt & Lee, 2002). Nonetheless, enrollment size still inflicted a relatively significant negative weight on progress scores of six

learning objectives: *creative capacities, broad liberal education, communication skills, critical analysis, diverse perspectives, and information literacy*. We, therefore, decided to retain enrollment size in the models. Of the two residual scores, we decided to retain  $D_N$  but not  $E_N$ .

*Self-efficacy* did not contribute strongly to most models, most likely because of its high correlation with *background preparation* ( $r = .75, p < .001$ ). However, it was the second most important variable in the model for developing *diverse perspectives* (Item 49). Because self-efficacy is a new variable the decision was made, therefore, to eliminate it from the adjustment, retain it in the instrument, and withhold judgment until more data are collected. Table 20 shows constants and unstandardized regression coefficients for the five predictor variables retained for producing adjusted scores on the 13 learning objectives and two overall summary measures in *Diagnostic Feedback* 2016.

Table 20

*Regression Analyses Predicting Student Ratings on Relevant Learning Objectives and Summary Measures with Course Motivation, Work Habits, Enrollment, Difficulty Residual, and Background Preparation*

Criterion	Constant	CM	WH	Enroll	$D_N$	BP	$R^2$	$N$	Grand Mean
23. Applications	0.927	.140	.333	-.001	-.007	.384	.306	1,801	4.13
24. Professional skills, viewpoints	0.783	.178	.414	-.001	-.032	.300	.313	1,135	4.18
25. Team skills	0.688	.076	.477	-.001	-.062	.326	.268	548	4.07
26. Creative capacities	1.729	.159	.333	-.010	-.028	.200	.237	308	4.15
27. Broad liberal education	1.731	.046	.388	-.006	-.007	.219	.181	432	4.06
28. Communication skills	1.455	.035	.375	-.010	-.110	.321	.230	1,007	4.03
31. Critical analysis	1.571	.058	.304	-.003	.083	.330	.226	1,067	4.10
49. Diverse perspectives	2.324	.114	.230	-.003	.019	.152	.124	554	4.07
50. Ethical reasoning	1.614	.140	.211	.000	.034	.293	.219	392	4.01
51. Civic engagement	0.915	.233	.280	-.001	-.047	.329	.291	463	4.06
52. Quantitative literacy	0.893	-.007	.469	-.004	.131	.368	.264	287	4.00
53. Information literacy	1.297	.010	.466	-.006	-.054	.268	.205	510	4.03
54. Understanding subject matter	1.316	.165	.260	-.001	.018	.337	.293	1,101	4.13
41. Excellence of teacher	1.539	.037	.294	-.001	.005	.405	.179	2,777	4.31
42. Excellence of course	0.427	.307	.285	-.001	-.009	.407	.336	2,777	4.09

Note. CM = Course Motivation, WH = Work Habits,  $D_N$  = Difficulty unrelated to the instructor, BP = Background Preparation. Classes with response rates less than 50% were excluded. Each model for learning objectives only included classes where the instructor rated the objective as relevant to the course.

A similar procedure was performed to determine the coefficients required for adjusting scores for *Learning Essentials* 2016. The regression models included Work Habits, Course Motivation, Background Preparation, and enrollment size as independent variables. Table 21 shows constants and unstandardized regression coefficients for computing adjusted scores in *Learning Essentials* 2016.

Table 21

*Regression Analyses Predicting Student Ratings on Relevant Learning Objectives and Summary Measures with Course Motivation, Work Habits, Enrollment, and Background Preparation*

Criterion	Constant	CM	WH	Enroll	BP	R <sup>2</sup>	N	Grand Mean
23. Applications	0.928	.140	.331	-.001	.386	.306	1,801	4.13
24. Professional skills, viewpoints	0.779	.175	.408	-.001	.310	.313	1,135	4.18
25. Team skills	0.687	.066	.469	-.001	.346	.266	548	4.07
26. Creative capacities	1.755	.157	.327	-.010	.202	.236	308	4.15
27. Broad liberal education	1.733	.046	.387	-.006	.220	.181	432	4.06
28. Communication skills	1.487	.020	.353	-.010	.351	.224	1,007	4.03
31. Critical analysis	1.542	.066	.325	-.002	.306	.222	1,067	4.10
49. Diverse perspectives	2.324	.115	.232	-.003	.148	.123	554	4.07
50. Ethical reasoning	1.616	.143	.217	.000	.284	.219	392	4.01
51. Civic engagement	0.929	.228	.271	-.001	.341	.290	463	4.06
52. Quantitative literacy	0.892	-.001	.510	-.004	.323	.257	287	4.00
53. Information literacy	1.311	.004	.457	-.006	.280	.204	510	4.03
54. Understanding subject matter	1.315	.167	.263	-.001	.332	.293	1,101	4.13
41. Excellence of teacher	1.539	.038	.295	-.001	.403	.179	2,777	4.31
42. Excellence of course	0.428	.306	.283	-.001	.410	.336	2,777	4.09

*Note.* CM = Course Motivation, WH = Work Habits, BP = Background Preparation. Classes with response rates less than 50% were excluded. Each model for learning objectives only included classes where the instructor rated the objective as relevant to the course.

We use the data from a fictional course CS 205 with an enrollment of 98 students as an example to illustrate how adjusted scores are calculated for *Diagnostic Feedback* 2016. The students in CS 205 report the following average ratings:

*Stimulating intellectual effort* = 3.24

Amount of coursework = 3.98

Difficulty of subject matter = 4.02

Course Motivation = 3.32

Work Habits = 4.01

Background Preparation = 3.85

Progress on objective *Applications* = 4.51

We first calculate the predicted value of *Difficulty of subject matter* using Equation 1 on p. 35:

$$.004 \times 3.24 + .773 \times 3.98 + .691 = 3.7805.$$

Using Equation 2 on p. 36, the residual of difficulty  $D_N$  is:

$$4.02 - 3.7805 = 0.2395.$$

The Formula for adjusting scores for progress on objectives and summary measures is as follows:

$$\text{Adjusted Score} = \text{Grand Mean} + (\text{Obtained Mean} - \text{Predicted Mean}) \times (1 + R^2)$$

To calculate the adjusted score for student mean ratings on the learning objective

*Applications*, we first compute the predicted mean using the coefficients in Table 20:

$$\text{Predicted mean} = 0.927 + .140 \times 3.32 + .333 \times 4.01 + (-.001) \times 98 + (-.007) \times 0.2395 + .384 \times 3.85 = 4.1058535.$$

Next, we compute the adjusted mean, using the Adjusted Score formula above.

Referring to Table 20, the Grand Mean for *Applications* = 4.13 and  $R^2 = .306$ , so:

$$\text{Adjusted Score} = 4.13 + (4.51 - 4.1058535) \times (1 + .306) = 4.657815329$$

### Reliabilities

Another important criterion in selecting items for inclusion in IDEA2 is class-level reliability, a measure of the consistency among student raters taking the same course from the same instructor. Ratings are less helpful in decision making if students vary substantially in their perceptions of progress on learning objectives, the quality of the course, and the effectiveness of the instructor.

To measure reliability at the class level we applied the procedures described in James, Demaree, and Wolf (1984, p. 87) for computing the *within-group interrater reliability coefficient*. Consistent with Hoyt et al. (1999) we performed the analyses on classes with 15-34 students enrolled, as described below. The average number of students responding in those classes was 23. We applied the following single-item interrater agreement formula to compute the reliability coefficients for existing and proposed items found in Table 22:

$$r_{WG(I)} = 1 - (S_{X_j}^2 / \sigma_{EU}^2)$$

where  $r_{WG(I)}$  is the within-group interrater reliability for a group of  $K$  judges on a single item  $X_j$ , and  $S_{X_j}^2$  is the observed variance of  $X_j$ ,  $\sigma_{EU}^2$  is the variance of  $X_j$  that would be expected if all judgments solely resulted from random measurement error. Thus  $\sigma_{EU}^2 = (A^2 - 1)/12$  where  $A$  corresponds to the number of alternatives in the response scale for  $X_j$ , which is presumed to vary from 1 to  $A$ .

Next, we computed the standard error of measurement on each item, applying the following formula:

$$SEM = S \sqrt{1 - r_{xx}}$$

where  $S$  refers to the standard deviation of the item and  $r_{xx}$  denotes the single-item reliability. SEM provides an estimate of the amount of error that likely would be associated with the obtained mean score on an individual item. Table 22 contains the SEM for existing and proposed items.

All reliability coefficients were at or above .80 and all SEM were below .3 with the exception of one item: *Formed "teams" or "discussion groups" to facilitate learning*. In spite of

this item's relatively lower reliability it will remain in a modified form—*Formed teams or groups to facilitate learning*—because of its importance in measuring the frequency of collaborative learning. We hypothesize that its reliability might increase with the change in wording because students would not be restricted to considering only *discussion* groups.

Table 22

*Within-group Interrater Reliability Coefficients and Standard Errors of Measurement of Items on the Pilot Diagnostic Feedback Instrument*

Item	<i>M</i>	<i>SD</i>	$r_{wg}(I)$	<i>SEM</i>
1. Displayed personal interest in students	4.40	0.46	.89	.15
2. Helped students answer own questions	4.26	0.50	.88	.17
3. Scheduled work helpfully	4.35	0.45	.90	.15
4. Demonstrated importance of subject	4.42	0.44	.91	.13
5. Formed teams, discussion groups	3.98	0.74	.73	.38
6. Made clear how topics fit	4.37	0.47	.89	.16
7. Explained criticisms	4.16	0.51	.87	.19
8. Stimulated intellectual effort	4.16	0.51	.87	.18
9. Encouraged use of multiple resources	4.11	0.54	.85	.21
10. Explained clearly	4.26	0.57	.84	.23
11. Related to real life	4.36	0.49	.88	.17
12. Tests covered important points	4.35	0.44	.90	.14
13. Introduced stimulating ideas	4.25	0.52	.87	.19
14. Involved students in hands on activities	4.04	0.63	.80	.28
15. Inspired students to set high goals	4.07	0.55	.85	.21
16. Asked students to share experiences	4.06	0.64	.80	.29
17. Provided timely feedback	4.24	0.56	.84	.22
18. Asked students to help each other	4.08	0.54	.85	.21
19. Assessments required creativity	4.18	0.52	.86	.19
20. Encouraged student/faculty contact	4.15	0.53	.86	.20
21. Factual knowledge	4.10	0.46	.90	.15
22. Principles and theories	4.09	0.46	.89	.15
23. Applications	4.12	0.48	.89	.16
24. Professional skills, viewpoints	4.09	0.48	.88	.17
25. Team skills	3.75	0.62	.81	.27
26. Creative capacities	3.70	0.62	.81	.27
27. Broad liberal education	3.82	0.56	.84	.22
28. Communication skills	3.79	0.62	.81	.27
29. Find, use resources	3.92	0.50	.88	.18
30. Values development	3.84	0.57	.84	.23
31. Critical analysis	3.97	0.53	.86	.20
32. Interest in learning	3.97	0.51	.87	.18

33. Amount of reading	3.26	0.63	.80	.28
34. Amount of other work	3.37	0.46	.89	.15
35. Difficulty of subject matter	3.38	0.51	.87	.18
36. Strong desire to take the course	3.66	0.63	.80	.28
37. Worked harder on this course than most	3.64	0.49	.88	.17
38. Wanted this instructor	3.64	0.63	.80	.28
39. Wanted course regardless of instructor	3.37	0.49	.88	.17
40. Increase positive attitude toward field	3.97	0.57	.84	.23
41. Excellent instructor	4.31	0.57	.84	.23
42. Excellent course	4.09	0.57	.84	.23
43. Usually work hard on academic work	3.89	0.29	.96	.06
44. Used variety of evaluation methods	4.08	0.48	.88	.16
45. Expected students to take responsibility	4.37	0.32	.95	.07
46. High achievement standards	4.26	0.38	.93	.10
47. Used educational technology	4.24	0.49	.88	.17
49. Diverse perspectives	3.76	0.62	.81	.27
50. Ethical reasoning	3.76	0.60	.82	.26
51. Civic engagement	3.87	0.56	.84	.22
52. Quantitative literacy	3.70	0.58	.83	.24
53. Information literacy	3.92	0.51	.87	.18
54. Understanding subject matter	4.13	0.46	.89	.15
55. Helped interpret subject matter	3.99	0.61	.82	.26
56. Encouraged self-reflection	4.21	0.50	.87	.18
57. Created service opportunities	3.99	0.57	.84	.23
58. Provided meaningful feedback	4.14	0.56	.84	.23
59. Self-efficacy	3.95	0.38	.93	.10
60. Background preparation	3.82	0.45	.90	.14
61. Amount of coursework	3.48	0.45	.90	.14

Note.  $N = 2,426$ .

### Updated Student Ratings of Instruction Instruments

#### Diagnostic Feedback

Based on the analyses described in this report, the items to be included in the revised *Diagnostic Feedback* instrument are as follows.

#### Teaching Methods

1. Found ways to help students answer their own questions
2. Helped students to interpret subject matter from diverse perspectives (e.g., different cultures, religions, genders, political views)
3. Encouraged students to reflect on and evaluate what they have learned

4. Demonstrated the importance and significance of the subject matter
5. Formed teams or groups to facilitate learning
6. Made it clear how each topic fit into the course
7. Provided meaningful feedback on students' academic performance
8. Stimulated students to intellectual effort beyond that required by most courses
9. Encouraged students to use multiple resources (e.g., Internet, library holdings, outside experts) to improve understanding
10. Explained course material clearly and concisely
11. Related course material to real life situations
12. Created opportunities for students to apply course content outside the classroom
13. Introduced stimulating ideas about the subject
14. Involved students in hands-on projects such as research, case studies, or real life activities
15. Inspired students to set and achieve goals which really challenged them
16. Asked students to share ideas and experiences with others whose backgrounds and viewpoints differ from their own
17. Asked students to help each other understand ideas or concepts
18. Gave projects, tests, or assignments that required original or creative thinking
19. Encouraged student-faculty interaction outside of class (e.g., office visits, phone calls, email)

#### *Learning Objectives*

20. Gaining a basic understanding of the subject (e.g., factual knowledge, methods, principles, generalizations, theories)
21. Developing knowledge and understanding of diverse perspectives, global awareness, or other cultures
22. Learning to *apply* course material (to improve thinking, problem solving, and decisions)
23. Developing specific skills, competencies, and points of view needed by professionals in the field most closely related to this course
24. Acquiring skills in working with others as a member of a team
25. Developing creative capacities (inventing; designing; writing; performing in art, music, drama, etc.)
26. Gaining a broader understanding and appreciation of intellectual/cultural activity (music, science, literature, etc.)
27. Developing skill in expressing myself orally or in writing
28. Learning how to find, evaluate, and use resources to explore a topic in depth
29. Developing ethical reasoning and/or ethical decision making
30. Learning to *analyze* and *critically evaluate* ideas, arguments, and points of view
31. Learning to apply knowledge and skills to benefit others or serve the public good
32. Learning appropriate methods for collecting, analyzing, and interpreting numerical information

*Course Characteristics*

- 33. Amount of coursework
- 34. Difficulty of subject matter

*Student Characteristics*

- 35. As a rule, I put forth more effort than other students on academic work.
- 36. I really wanted to take this course regardless of who taught it.
- 37. When this course began I believed I could master its content.
- 38. My background prepared me well for this course's requirements.

*Overall Summary Measures*

- 39. Overall, I rate this instructor an excellent teacher.
- 40. Overall, I rate this course as excellent.

**Learning Essentials***Learning Objectives*

- 1. Gaining a basic understanding of the subject (e.g., factual knowledge, methods, principles, generalizations, theories)
- 2. Developing knowledge and understanding of diverse perspectives, global awareness, or other cultures
- 3. Learning to *apply* course material (to improve thinking, problem solving, and decisions)
- 4. Developing specific skills, competencies, and points of view needed by professionals in the field most closely related to this course
- 5. Acquiring skills in working with others as a member of a team
- 6. Developing creative capacities (inventing; designing; writing; performing in art, music, drama, etc.)
- 7. Gaining a broader understanding and appreciation of intellectual/cultural activity (music, science, literature, etc.)
- 8. Developing skill in expressing myself orally or in writing
- 9. Learning how to find, evaluate, and use resources to explore a topic in depth
- 10. Developing ethical reasoning and/or ethical decision making
- 11. Learning to *analyze* and *critically evaluate* ideas, arguments, and points of view
- 12. Learning to apply knowledge and skills to benefit others or serve the public good
- 13. Learning appropriate methods for collecting, analyzing, and interpreting numerical information

*Student Characteristics*

- 14. As a rule, I put forth more effort than other students on academic work.
- 15. I really wanted to take this course regardless of who taught it.
- 16. My background prepared me well for this course's requirements.

### Overall Summary Measures

17. Overall, I rate this instructor an excellent teacher.
18. Overall, I rate this course as excellent.

### Suggested Actions for Teaching Methods

One of the distinguishing characteristics of the DF instrument is the formative feedback provided to instructors. We accomplish this by comparing the instructor's average class ratings of how frequently students perceived each of 19 teaching methods with those of similar class sizes and levels of student motivation. To obtain the average rating for each teaching method in comparable courses, we grouped classes based on the number of students enrolled (Small = 10-14, Medium = 15-34, Large = 35-49, and Very Large = 50 or greater) by average class ratings on *course motivation* ("I really wanted to take this course regardless of who taught it."). Five levels of course motivation were defined as High (the top 10 percent), High Average (the next 20 percent), Average (The middle 40 percent), Low Average (the next 20 percent), and Low (the bottom 10 percent). Table 23 presents the average ratings of the 15 existing teaching methods in research datasets from 2011 to 2015 and the pilot study, depending on class size and student motivation.

Table 23

*Average Student Ratings for Existing Teaching Method Items in Diagnostic Feedback by Class Size and Level of Course Motivation*

Course motivation level	Class size			
	Small	Medium	Large	Very Large
2. Helped students answer own questions				
Low (lowest 10%)	4.10	4.05	3.96	3.87
Low average (next 20%)	4.29	4.23	4.15	4.08
Average (Middle 40%)	4.41	4.36	4.29	4.22
High average (next 20%)	4.50	4.46	4.41	4.35
High (highest 10%)	4.59	4.56	4.54	4.51
4. Demonstrated significance				
Low	4.24	4.20	4.18	4.13
Low average	4.43	4.39	4.36	4.34
Average	4.55	4.52	4.50	4.47
High average	4.64	4.62	4.62	4.59
High	4.72	4.70	4.71	4.67
5. Formed teams				
Low	3.76	3.73	3.38	3.23
Low average	3.92	3.88	3.60	3.41
Average	4.06	4.04	3.80	3.57
High average	4.17	4.16	3.96	3.74
High	4.28	4.29	4.21	4.11
6. Made clear how topic fits				
Low	4.13	4.11	4.08	4.03
Low average	4.34	4.31	4.28	4.24

Average	4.47	4.45	4.43	4.39
High average	4.57	4.56	4.55	4.52
High	4.66	4.65	4.64	4.60
8. Stimulated intellectual effort				
Low	3.96	3.87	3.76	3.70
Low average	4.14	4.06	3.98	3.95
Average	4.28	4.21	4.15	4.11
High average	4.38	4.33	4.30	4.29
High	4.48	4.45	4.48	4.47
9. Encouraged using multiple resources				
Low	3.94	3.87	3.60	3.55
Low average	4.09	4.01	3.81	3.76
Average	4.20	4.14	4.00	3.91
High average	4.29	4.24	4.16	4.08
High	4.37	4.35	4.36	4.37
10. Explained clearly				
Low	4.00	4.01	3.95	3.89
Low average	4.22	4.21	4.17	4.14
Average	4.35	4.35	4.31	4.27
High average	4.47	4.45	4.43	4.40
High	4.57	4.56	4.53	4.50
11. Related to real life				
Low	4.08	4.07	4.10	4.09
Low average	4.32	4.29	4.32	4.30
Average	4.46	4.45	4.48	4.45
High average	4.57	4.56	4.61	4.58
High	4.65	4.65	4.69	4.66
13. Introduced stimulating ideas				
Low	3.94	3.91	3.86	3.82
Low average	4.21	4.15	4.11	4.07
Average	4.37	4.33	4.30	4.26
High average	4.49	4.46	4.45	4.41
High	4.59	4.57	4.55	4.52
14. Involved in hands on				
Low	3.80	3.65	3.38	3.28
Low average	4.06	3.87	3.64	3.48
Average	4.24	4.11	3.89	3.69
High average	4.38	4.29	4.11	3.93
High	4.51	4.44	4.35	4.29
15. Inspired ambitious goals				
Low	3.83	3.73	3.55	3.49
Low average	4.05	3.94	3.80	3.76
Average	4.22	4.12	4.01	3.93
High average	4.36	4.28	4.20	4.14

High	4.50	4.45	4.43	4.41
16. Asked diverse students to share ideas				
Low	3.77	3.71	3.44	3.32
Low average	4.00	3.92	3.72	3.60
Average	4.16	4.10	3.94	3.74
High average	4.28	4.22	4.10	3.90
High	4.39	4.36	4.31	4.26
18. Asked students to help others				
Low	3.88	3.81	3.64	3.56
Low average	4.08	4.01	3.86	3.80
Average	4.23	4.16	4.04	3.94
High average	4.34	4.30	4.20	4.11
High	4.46	4.43	4.41	4.37
19. Required originality				
Low	3.95	3.91	3.64	3.54
Low average	4.16	4.09	3.88	3.73
Average	4.30	4.24	4.06	3.88
High average	4.41	4.34	4.17	4.00
High	4.49	4.44	4.31	4.24
20. Encouraged out-of-class contact				
Low	4.00	3.94	3.87	3.83
Low average	4.15	4.11	4.07	4.04
Average	4.27	4.24	4.21	4.18
High average	4.37	4.34	4.35	4.31
High	4.46	4.45	4.49	4.46
<i>N</i>				
Low	4,297	22,323	1,982	743
Low average	8,965	42,720	4,855	1,531
Average	21,339	84,582	8,916	3,135
High average	13,484	39,552	3,260	1,537
High	9,312	18,461	1,562	959

*Note.* Classes with response rates lower than 75% in the 2011-2015 research datasets and those with response rates lower than 50% in the pilot dataset were excluded. The cutoff values for motivation levels were 2.78, 3.22, 3.81, and 4.20.

Given the limited number of classes in the pilot dataset, we considered it unwise to follow the above-mentioned procedure for the four new teaching methods. We instead employed the same two-step strategy as Hoyt et al. (1999) used to obtain estimates of class size-by-motivation level means. First, we computed the means in the pilot dataset for classes grouped by number of students enrolled (Small = 10-14, Medium = 15-34, Large = 35-49, and Very Large = 50 and more). Next, we grouped medium-size classes by average level of motivation, using the same criteria applied to the 15 existing teaching methods. We then calculated the means of the four new teaching methods in medium-sized classes, depending on students' average level of course motivation. We assumed the differences between motivation levels found in medium-size classes would be similar to those in the other-size classes. Based on this assumption, we

estimated means for each of the 20 class size-by-motivation-level groups. Table 24 presents statistics used to estimate average ratings of proposed teaching methods for suggested actions.

Table 24

*Average Student Ratings for Proposed Teaching Methods by Class Size and Average Student Ratings for Proposed Teaching Methods by Class Size and Level of Course Motivation in Medium-Sized Classes*

Item	Class size			Motivation level in Medium-sized classes					
	Small	Medium	Large	Very Large	Low	Average	Average	High	High
55. Helped interpret subject matter	4.03	4.01	3.92	3.75	3.81	3.93	4.01	4.12	4.17
56. Encouraged self-reflection	4.33	4.22	4.14	4.05	4.01	4.14	4.21	4.34	4.42
57. Created service opportunities	4.08	4.01	3.92	3.76	3.72	3.90	4.02	4.14	4.22
58. Provided meaningful feedback	4.25	4.17	4.04	3.82	3.98	4.07	4.17	4.25	4.35
<i>N</i>	367	2,099	206	105	228	419	836	388	228

The estimation process is explained as follows, using Item 56 as an example. As shown in Table 24, the average ratings of Item 56 in small, medium, large, and very large classes were 4.33, 4.22, 4.14 and 4.05 respectively. To get estimates of ratings in small classes by motivation levels, we applied the difference between ratings in small and medium classes to those in medium classes by motivation levels. The resulting estimates for small classes by motivation level were  $4.01 + (4.33 - 4.22) = 4.12$ ,  $4.14 + (4.33 - 4.22) = 4.25$ ,  $4.21 + (4.33 - 4.22) = 4.32$ ,  $4.34 + (4.33 - 4.22) = 4.45$ , and  $4.42 + (4.33 - 4.22) = 4.53$ . Table 25 displays the estimated average ratings of proposed teaching methods in all class sizes.

Table 25

*Estimated Average Student Ratings for Proposed Teaching Method Items in Diagnostic Feedback by Class Size and Level of Course Motivation*

Course motivation level	Class size			
	Small	Medium	Large	Very Large
55. Helped interpret subject matter				
Low (lowest 10%)	3.83	3.81	3.73	3.55
Low average (next 20%)	3.95	3.93	3.84	3.67
Average (Middle 40%)	4.02	4.01	3.92	3.75
High average (next 20%)	4.13	4.12	4.03	3.85
High (highest 10%)	4.18	4.17	4.08	3.90
56. Encouraged self-reflection				
Low	4.12	4.01	3.93	3.84
Low average	4.25	4.14	4.06	3.97
Average	4.32	4.21	4.13	4.04
High average	4.45	4.34	4.25	4.16
High	4.53	4.42	4.33	4.25
57. Created service opportunities				
Low	4.12	4.01	3.93	3.84

Low average	4.25	4.14	4.06	3.97
Average	4.32	4.21	4.13	4.04
High average	4.45	4.34	4.25	4.16
High	4.53	4.42	4.33	4.25
58. Provided meaningful feedback				
Low	4.07	3.98	3.85	3.64
Low average	4.16	4.07	3.95	3.73
Average	4.26	4.17	4.04	3.83
High average	4.34	4.25	4.13	3.91
High	4.44	4.35	4.22	4.01

*Note.* Classes with response rates lower than 50% in the pilot dataset were excluded. The cutoff values for motivation levels were 2.76, 3.15, 3.67, and 4.00.

### References

- Association of American Colleges and Universities. (2014). VALUE: Valid Assessment of Learning in Undergraduate Education. Retrieved from: [http://www.aacu.org/value/rubrics/index\\_p.cfm](http://www.aacu.org/value/rubrics/index_p.cfm)
- Benton, S. L., Duchon, D., & Pallett, W. H. (2011). Validity of self-reported student ratings of instruction. *Assessment & Evaluation in Higher Education*, 38, 377-389.
- Benton, S. L., Li, D., Brown, R., Guo, M., & Sullivan, P. (2015). *IDEA Technical Report No. 18: Revising the IDEA Student Ratings of Instruction System*. Manhattan, KS: The IDEA Center. [http://ideaedu.org/wp-content/uploads/2015/12/Technical\\_report\\_18.pdf](http://ideaedu.org/wp-content/uploads/2015/12/Technical_report_18.pdf)
- Hoeting, J. A., Madigan, D., Raftery, A. E., & 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. <http://ideaedu.org/wp-content/uploads/2014/11/techreport-01.pdf>
- 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. <http://ideaedu.org/wp-content/uploads/2014/11/techreport-11.pdf>
- Hoyt, D. P., & Lee, E. (2002). *IDEA Technical Report No. 12: Basic data for the revised IDEA system*. Kansas State University, Manhattan, KS: The IDEA Center. <http://ideaedu.org/wp-content/uploads/2014/11/techreport-12.pdf>
- James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, 69, 85-98.

### Appendix A: 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 <sup>1</sup>
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 <i>apply</i> course material (to improve thinking, problem solving, and decisions)	3. Learning to <i>apply</i> 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	4. Developing specific skills, competencies, and points of view needed by professionals in the field most	No change	

<sup>1</sup> 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

most closely related to this course	closely related to this course		
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 parenthetical 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 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 <i>evaluating</i> 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 <i>analyze</i> and <i>critically</i> 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 B: 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 <sup>2</sup>
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

<sup>2</sup> Changes are based on extensive research and feedback from experts panels and focus groups

			recommended it. 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.

### Appendix C: IDEA Alignment with NSSE Engagement Indicators

Themes	NSSE Engagement Indicators	Teaching Methods: original NSSE	New TMs based on IDEA categories	New Learning Objectives
<b>Academic Challenges</b>	<p><b>Higher-Order Learning:</b></p> <p>Challenging intellectual and creative work is central to student learning and collegiate quality. Colleges and universities promote high levels of student achievement by calling on students to engage in complex cognitive tasks requiring more than mere memorization of facts. This Engagement Indicator captures how much students' coursework emphasizes challenging cognitive tasks such as application, analysis, judgment, and synthesis.</p>	<p>9. Encouraged students to use multiple resources (e.g., Internet, library holdings, outside experts to improve understanding)</p> <p>14. Involved students in hands-on projects such as research, case studies, or real life activities</p> <p>18. Gave projects, tests, or assignments that required original or creative thinking</p>	<p><i>Active Learning</i></p> <p>9. Encouraged students to use multiple resources (e.g., Internet, library holdings, outside experts) to improve understanding</p> <p>14. Involved students in hands-on projects such as research, case studies, or real life activities</p> <p>18. Gave projects, tests, or assignments that required original or creative thinking</p>	<ul style="list-style-type: none"> <li>• Learning to <i>apply</i> course material (to improve thinking, problem solving, and decisions)</li> <li>• Developing creative capacities (inventing; designing; writing; performing in art, music, drama, etc.)</li> <li>• Developing skill in expressing oneself orally and in writing</li> <li>• Learning how to find, evaluate, and use resources to explore a topic in depth</li> <li>• Learning to <i>analyze</i> and <i>critically evaluate</i> ideas, arguments, and points of view</li> <li>• Learning appropriate methods for collecting, analyzing, and interpreting numerical information</li> </ul>
	<p><b>Reflective &amp; Integrative Learning:</b></p> <p>Personally connecting with course material requires students to relate their understandings and experiences to the content at hand. Instructors emphasizing reflective and integrative learning</p>	<p>2. Helped students to interpret subject matter from diverse perspectives (e.g., different cultures, religions, genders, political views)</p> <p>3. Encouraged students to reflect</p>	<p><i>Reflective and Integrative Learning</i></p> <p>2. Helped students to interpret subject matter from diverse perspectives (e.g., different cultures, religions, genders,</p>	<ul style="list-style-type: none"> <li>• Developing specific skills, competencies, and points of view needed by professionals in the field most closely related to this course</li> </ul>

	<p>motivate students to make connections between their learning and the world around them, reexamining their own beliefs and considering issues and ideas from others' perspectives.</p>	<p>on and evaluate what they have learned</p> <p>8. <i>Stimulated students to intellectual effort beyond that required by most courses</i></p> <p>11. Related course material to real life situations</p> <p>12. Created opportunities for students to apply course content outside the classroom</p>	<p>political views)</p> <p>3. Encouraged students to reflect on and evaluate what they have learned</p> <p><b>7. <i>Provided meaningful feedback on students' academic performance</i></b></p> <p>8. Stimulated students to intellectual effort beyond that required by most courses</p> <p>11. Related course material to real life situations</p> <p>12. Created opportunities for students to apply course content outside the classroom</p>	<ul style="list-style-type: none"> <li>• Gaining a broader understanding and appreciation of intellectual /cultural activity</li> <li>• Developing ethical reasoning and/or ethical decision making</li> <li>• Learning to apply knowledge and skills to benefit others and serve the public good</li> <li>• Developing knowledge and understanding of diverse perspective, global awareness, or other cultures</li> <li>• Learning appropriate methods for collecting, analyzing, and interpreting numerical information</li> </ul>
<b>Learning with Peers</b>	<p><b>Collaborative Learning:</b></p> <p>Collaborating with peers in solving problems or mastering difficult material deepens understanding and prepares students to deal with the messy, unscripted problems they encounter during and after college. Working on group projects, asking others for help with difficult material or explaining it to others, and working through course material in preparation for exams all represent collaborative learning activities.</p>	<p>5. Formed teams or groups to facilitate learning</p> <p>16. <i>Asked students to share ideas and experiences with others whose backgrounds and viewpoints different from their own</i></p> <p>17. Asked students to help each other understand ideas or concepts</p>	<p><i>Collaborative Learning</i></p> <p>5. Formed teams or groups to facilitate learning</p> <p>16. Asked students to share ideas and experiences with others whose backgrounds and viewpoints differ from their own</p> <p>17. Asked students to help each other understand ideas or concepts</p>	<ul style="list-style-type: none"> <li>• Acquiring skills in working with others as a member of a team</li> </ul>
	<p><b>Discussion with Diverse Others:</b></p>	<p><b>Teaching Methods:</b></p> <p>16. Asked students to share ideas</p>		<ul style="list-style-type: none"> <li>• Developing knowledge and understanding of</li> </ul>

	Colleges and universities afford students new opportunities to interact with and learn from others with different backgrounds and life experiences. Interactions across difference, both inside and outside the classroom, confer educational benefits and prepare students for personal and civic participation in a diverse and interdependent world	and experiences with others whose backgrounds and viewpoints different from their own	diverse perspective, global awareness, or other cultures	
<b>Experiences with Faculty</b>	<p>Student-Faculty Interaction:</p> <p>Interactions with faculty can positively influence the cognitive growth, development, and persistence of college students. Through their formal and informal roles as teachers, advisors, and mentors, faculty members model intellectual work, promote mastery of knowledge and skills, and help students make connections between their studies and their future plans</p> <p>Effective Teaching Practices:</p> <p>Student learning is heavily dependent on effective teaching. Organized instruction, clear explanations, illustrative examples, and effective feedback on student work all represent aspects of teaching effectiveness that promote student comprehension and learning.</p>	<p>Teaching Methods:</p> <p>1. Found ways to help students answer their own questions</p> <p>19. Encouraged student-faculty interaction outside of class (office visits, phone calls, email etc.)</p> <p>4. Demonstrated the importance and significance of the subject matter</p> <p>6. Made it clear how each topic fit into the course</p> <p>10. Explained course material clearly and concisely</p> <p>13. Introduced stimulating ideas about the subject</p> <p>15. Inspired students to set and achieve goals which really challenged them</p> <p><b>7. Provided meaningful feedback on students' academic performance</b></p>	<p><i>Teaching Essentials</i></p> <p><i>1. Found ways to help students answer their own questions</i></p> <p>4. Demonstrated the importance and significance of the subject matter</p> <p>6. Made it clear how each topic fit into the course</p> <p>10. Explained course material clearly and concisely</p> <p>13. Introduced stimulating ideas about the subject</p> <p>15. Inspired students to set and achieve goals which really challenged them</p> <p>19. Encouraged student-faculty interaction outside of class (office visits, phone calls, e-mail, etc.)</p>	<p>NA- the above themes are related to learning and this one is about experiences with faculty</p>

### Appendix D: Bayesian Model Averaging on 13 Learning Objectives in Medium Classes (15-34)

#### Objective 3

Model Size	R <sup>2</sup>	TM1	TM2	TM3	TM4	TM5	TM6	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM18	TM19	TM20	TM55	TM56	TM57	TM58	
Full	.85							.11			.07	.08			.08						.06	.05		
2 vars	.79											.14			.27									
3 vars	.82											.12			.18						.13			
4 vars	.84							.17			.08	.09											.10	
5 vars	.84							.10			.08	.08			.10						.08			
6 vars	.85							.11			.07	.08			.08						.06	.05		

#### Objective 4

Model Size	R <sup>2</sup>	TM1	TM2	TM3	TM4	TM5	TM6	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM18	TM19	TM20	TM55	TM56	TM57	TM58	
Full	.81						.13				.06				.15								.06	
2 vars	.79						.18								.21									
3 vars	.81						.14				.08				.19									
4 vars	.81						.13				.06				.15									.06

#### Objective 5

Model Size	R <sup>2</sup>	TM1	TM2	TM3	TM4	TM5	TM6	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM18	TM19	TM20	TM55	TM56	TM57	TM58	
Full	.68					.17									.21									
2 vars	.68					.17									.21									

#### Objective 6

Model Size	R <sup>2</sup>	TM1	TM2	TM3	TM4	TM5	TM6	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM18	TM19	TM20	TM55	TM56	TM57	TM58	
Full	.63										-.15	-.09		.37			.15							
2 vars	.57										-.15			.41										
3 vars	.61										-.17			.35			.11							
4 vars	.63										-.15	-.09		.37			.15							

#### Objective 7

Model Size	R <sup>2</sup>	TM1	TM2	TM3	TM4	TM5	TM6	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM18	TM19	TM20	TM55	TM56	TM57	TM58	
Full	.78						.13						.19							.11				
2 vars	.76						.13						.28											
3 vars	.78						.13						.19							.11				



## Objective 52

Model Size	R <sup>2</sup>	TM1	TM2	TM3	TM4	TM5	TM6	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM18	TM19	TM20	TM55	TM56	TM57	TM58	
Full	.68											.64			-.24			.40						
1 var	.60											.78												
2 vars	.65											.47						.38						
3 vars	.68											.64			-.24			.40						

## Objective 53

Model Size	R <sup>2</sup>	TM1	TM2	TM3	TM4	TM5	TM6	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM18	TM19	TM20	TM55	TM56	TM57	TM58
Full	.72							.29	.24														.41
1 var	.63																						.79
2 vars	.70							.41															.48
3 vars	.72							.29	.24														.41

## Objective 54

Model Size	R <sup>2</sup>	TM1	TM2	TM3	TM4	TM5	TM6	TM8	TM9	TM10	TM11	TM12	TM13	TM14	TM15	TM16	TM18	TM19	TM20	TM55	TM56	TM57	TM58
Full	.76							.35			.21	.18				-.21							.41
1 var	.65							.80															
2 vars	.71							.50														.40	
3 vars	.74							.39				.23										.34	
4 vars	.75							.43			.27					-.22						.44	
5 vars	.76							.35			.21	.18				-.21						.41	