This paper describes the development of a student rating of teaching effectiveness (SRTE) instrument by Saint Leo College (Florida). The instrument was based on Cashin's (1989) seven dimensional model of college teaching: subject matter mastery, curriculum development, course design, delivery of instruction, assessment of student learning, availability to students, and administrative requirements. A 22-item index, based on the student-perspective dimension of Cashin's model, was developed; it included 3 global items, 5 general concept items on course design, and 14 general concept items over several aspects of instruction. In the fall of 1993 the index was administered to 146 classes under standard conditions. Factors analysis of the 19 nonglobal items indicated relatively high correlations between all variables. The high correlations presented a problem for the SRTE instrument, in that it is most unlikely that teaching is a unidimensional construct. Thus, an alternative explanation offered is that a strong "halo" effect is operating and student differentiation among the dimensions is obscured. An appendix provides instructional dimensions, student-perspective dimensional attributes, a copy of the SRTE index, and recommendations for using student ratings of teaching. (Contains 40 references.) (MDM)
Assessing Teaching Effectiveness in a Liberal Arts College: 
The Student Perspective

Charles D. Hale
Coordinator: Academic Programs, Assessment, & Evaluation
Department of Evening Classes
University of Georgia Center for Continuing Education
Athens GA 30602-3603

Charlene Herreid
Director, Institutional Research
Saint Leo College
St. Leo FL 33574

Gordon Waugh
Research Scientist
Human Resources Research Organization (HumRRO)
Alexandria Va 22314
Assessing Teaching Effectiveness: The Student Perspective

Student ratings of teaching effectiveness have been the subject of much empirical research and debate (Abrami, 1985, 1989a, 1989b; Abrami & d’Apollonia, 1991; Arreola, 1986; Cashin, 1988, 1989, 1990a; Cashin & Downey 1992; Cashin, Downey, & Sixbury, 1994; Centra, 1979; Cohen, 1981; Marsh, 1982, 1984, 1994). Additionally, Student ratings of teaching effectiveness have gained an increasingly important role in higher education teaching effectiveness assessment and faculty development systems (Arreola, 1995; Seldin, 1993; Ory & Parker, 1989; Theall & Franklin, 1990). According to Marsh (1987, p. 259) student ratings of teaching effectiveness (SRTE) are used for five purposes

1. Diagnostic feedback to faculty about the effectiveness of their teaching that will be useful for the improvement of teaching;
2. A measure of teaching effectiveness to be used in administrative decision-making;
3. Information for students to use in the selection of courses and instructors;
4. A measure of the quality of the course, to be used in course improvement and curriculum development;
5. An outcome or process description for research on teaching.

Purposes one and two are widely followed in the American Academy; while, the remaining three are less followed. What is clear is that student ratings of teaching effectiveness are integral to post secondary faculty evaluation and development.

Despite their widespread use for the past 60 or so years, student ratings of instruction (i.e., the instructor and course) are subject to faculty and administrative misconceptions. According to Cohen (1990, p. 124), many faculty and administrators hold several misconceptions which include

a. Students are not qualified to make judgments about teaching competence.
b. Student ratings are popularity contests.
c. Students are unable to make accurate judgments until after they have been away from the course for several years.
d. Student ratings are unreliable.
e. Students ratings are invalid.
f. Students rate instructors on the basis of grades they receive.
g. Extraneous variables and conditions affect student ratings.

While a discussion of each of these "myths" is beyond the scope of the present paper, Marsh (1984, p. 707), after an extensive literature review, concluded

class-average student ratings are (a) multidimensional; (b) reliable and stable; (c) primarily a function of the instructor who teaches a course rather than the course that is taught; (d)
relatively valid against a variety of indicators of effective teaching; (e) relatively unaffected by
a variety of variables hypothesized as potential biases; and (f) seen to be useful by faculty as
feedback about their teaching, by students for use in course selection, and by administrators
for use in personnel decisions.

For additional information, the interested reader is invited to consult Feldman (1989a, 1989b);
Marsh (1987, 1991); and Murry, Ruston, and Paunonen (1990) for extensive discussions or Cashin

Creating an Assessment Instrument

In the spring of 1994, Saint Leo College determined that as part of its attempt to improve the
assessment of teaching at the college, a college-wide standardized questionnaire would be instituted
to collect information on student ratings of teacher effectiveness. This standardized instrument
would replace a "hodge-podge" of "homegrown" questionnaires of unknown reliability and validity
used by various departments throughout the college. The new instrument would be used for
personnel decision-making and to provide feedback to professors about student perceptions of
teaching effectiveness.

In searching for an instrument to use, we informally surveyed other similar institutions. A
bewildering array of tools was found to be in use. Most of these were not documented as to their
reliability or validity, and several institutions could not delineate the characteristics of effective
teaching their instruments were intended to assess. In explaining why there exist differences
between item pools on student ratings of teaching effectiveness (SRTE) forms, Abrami (1985)
observed

The general lack of a sophisticated theoretical rationale for describing effective college teaching
and selecting items for analysis may explain why item pools differ. Instead of relying on
theory to guide item selection, item pools have been generated by faculty and student
committees, through student descriptions of ideal professors or good teaching, [and] by
selecting items form other rating forms etc. (p. 216)

Abrami (1985) goes on to state, "[f]urther progress in measuring college teaching awaits attention
to developing and utilizing theories of instruction appropriate for higher education." We found this
to be the case.

Some commercial instruments were also examined. Although their item pools seemed to be
better documented and more firmly based on sound premises, the cost of using these rating systems (between .25 and .50 cents per answer sheet scored) was prohibitive as we expected approximately 50,000 SRTEs to be processed each fiscal year. Thus, we decided to produce and validate our own instrument to assess student ratings of teaching effectiveness.

Towards a Theory: Cashin's Model

Cashin's (1989) model of college teaching seemed a promising theoretical model. After carefully reviewing the work of Centra (1977, 1979) and Arreola (1986, 1989), Cashin (1989) advanced a seven dimensional model of college teaching, which includes (a) subject matter mastery, (b) curriculum development, (c) course design, (d) delivery of instruction, (e) assessment of instruction, (relabeled assessment of student learning (Cashin, personal communication 1995), (f) availability to students, and (g) administrative requirements. Cashin further argued that there are five principal perspectives from which teaching and learning should be assessed; these are (a) the instructor, himself or herself; (b) students; (c) peers, persons who are knowledgeable in the subject matter; (d) colleagues, persons who are knowledgeable about teaching but not the specific subject matter; and (e) the department head or dean. Cashin does acknowledge that other administrative personnel and/or an instructional consultant, if available, may have an interest in faculty evaluation.

Next, Cashin (1989) specifies which teaching dimensions each perspective is competent to assess. According to Cashin (1989), students are capable of assessing the delivery of instruction, assessment of student learning, availability, and administrative requirements (selected aspects) dimensions. Peers are competent to assess the subject matter mastery, curriculum development, course design, delivery of instruction, and assessment of student learning dimensions. Colleagues can accurately assess the delivery of instruction and assessment of student learning dimensions. The department head or dean can assess the curriculum development, course design, assessment of student learning, and administrative requirement dimensions. An instructional consultant, if available, can assess the course design, delivery of instruction, and assessment of student learning dimensions. Taken together, these perspectives yield data upon which judgments about an
instructor's teaching effectiveness can be made. Thus, we elected to base our SRTE on the student perspective dimension of Cashin's model.

**Global, General Concept, and Specific Items**

In drafting items to be included in our SRTE, we needed to determine whether to include global, general concept, or specific items. The Office of Instructional Resources (no date) at the University of Illinois at Urbana-Champaign, has outlined a three tier item type classification scheme: global, general concept, and specific. **Global** items are very general in wording and are intended to be comparative across a variety of disciplines and instructional contexts. **General concept** items "may best be described as 'indicator items'-that indicate a general area of strength or weakness" (p. 3) and are more likely to be used for administrative purposes. **Specific** items are those "which request reports of class activities or observations of instructor behaviors so they do not necessarily require summary judgments...[and] specific items may not be necessarily evaluative" (p.3). Murry (1983) labels these as low-inference items.

The authors state that there are two general criteria which influence the classification of items; they are (a) "how specific the item is in requesting student judgments or observations about a course and (b) the use to be made of the information" (p. 2). Item specificity is the most important of the two. Global and general concept items tend to have more administrative utility whereas specific items are mostly used by an instructor (p. 2). General concept items and specific items (often referred to as low-inference items) tend to have lower inference value than global items.

Global items are recommended for summative personnel decisions (Abrami, 1989b; Abrami & d’Apollonia, 1990; Braskamp & Ory, 1994; Cashin & Downey, 1992; Scriven, 1981). Since our instrument was intended to be used for both personnel decision-making and individual diagnostic purposes, both global and general concept item types were included.

**Interpreting Results**

With respect to the interpretation, McKeachie and Kaplan (1996) recommend that SRTE data be summarized by response option percentage distribution for each item. It is also common to
report item means and standard deviations. Cashin (1992) offers alternative interpretation strategies. If subtest scores, as advocated by Marsh, are used, summary data should be organized by subtest for the user's convenience.

Low-inference items are by definition of very limited generalizability. Hence, their interpretation is context bound and the information is most useful to the rated instructor for teaching improvement. If an instructional consultant is available, he or she can use such data for individual consultation directed towards improving teaching quality.

**Methodology**

**The Instrument**

A 22 item index, based on the student perspective dimension of Cashin's model, was organized into three sections: (a) three global items; (b) five general concept items on course design, etc.; and (c) 14 general concept items over several aspects of instruction. Both global and general concept items were employed (Cashin and Downey, 1992; Office of Instructional Resources, no date; Marsh, 1994) as the index was to provide data for personnel decision-making, across academic disciplines with similar instructional missions and methods and for individual diagnostic purposes. Response options for Items 1 through Item 8 were along a six point Likert style scale from very strongly disagree (1) to very strongly agree (6). The response continuum for items 4 to 22 was from never (1) to always (6).

Aside from the three global items, the index was designed to comprise four subtests: academic administration, delivery of instruction, availability to students, and assessment of student learning. The academic administration dimension was envisioned to address aspects of course design, content relevance to course objectives, and expectation clarity; thus, it was operationalized by items 4 to 8. The delivery of instruction dimension was expected to assess relevance of assignments and examinations to course content, instructor rapport with students, and feedback quality; defining items were 9 to 17 and 18. The availability to student dimension was defined by a single item (16), an obvious design flaw, as an entire dimension should never be defined by a single item. The
assessment of student learning dimension was intended to assess the influence of examinations,
assignments, teaching methods, and textbooks on stimulating student learning; items comprising
the dimension were 19, 20, 21, and 22.

In the fall of 1993, the index was administered in 146 classes under standard conditions with
faculty reading from a script and then exiting the room. Evaluations were collected by a designated
student and delivered to the college’s office of institutional research.

**Factor Analysis**

First, a principal axis extraction without rotation was conducted to determine the number of
factors to retain using squared multiple correlations to estimate the communalities. The 19 lower
inference items (4-22) from the Saint Leo College SRTE were examined via factor analysis (FA)
using the SAS statistical package. Global items (1-3) were excluded from factoring. Data from
1,786 evaluations in 146 courses were included in the analysis; 60 cases were deleted due to
missing data. Item means and standard deviations are presented in Table 1. Responses to all 19
items were negatively skewed (most responses were positive); nevertheless, the correlation matrix
revealed moderately high correlations between all variables (.50 to .80).

After deleting Item 22 due to a low communality value, a second principal axis FA without
rotation was run. The associated scree plot showed a break at four factors; however, a parallel
analysis indicated six. Given these data, a set of principal axis FA procedures, using the Harris-
Kaiser rotation with maximum oblique rotation, was run, retaining four, five and six factors. While
the five factor solution seemed to be the most meaningful, Items 9 and 10 loaded on an
uninterpretable factor and in an unstable fashion on other factors. Hence, Items 9 and 10 were
deleted.

Again, a principal axis FA without rotation was run with the associated scree plot indicating
the presence of three factors; however, a parallel analysis revealed the possible presence of five
factors. Again, the FA procedure with the same rotation was repeated, retaining three, four, and
five factors. The four factor solution was the most interpretable.
Results

Presented in Table 2 are the rotated factor loadings using .45 as the significance criterion.

<table>
<thead>
<tr>
<th>Item</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. This course contributed to professional and/or professional development.</td>
<td>4.899</td>
<td>1.438</td>
</tr>
<tr>
<td>2. The instructor's teaching was effective helping me learn.</td>
<td>4.911</td>
<td>1.399</td>
</tr>
<tr>
<td>3. The instructor showed respect for students.</td>
<td>5.004</td>
<td>1.545</td>
</tr>
<tr>
<td>4. Course learning objectives were clearly stated and explained.</td>
<td>5.026</td>
<td>1.254</td>
</tr>
<tr>
<td>5. Course content was related to learning objectives.</td>
<td>5.090</td>
<td>1.143</td>
</tr>
<tr>
<td>6. Course organization was logical and understandable.</td>
<td>4.908</td>
<td>1.438</td>
</tr>
<tr>
<td>7. The syllabus clearly explained course organization and expectations.</td>
<td>5.046</td>
<td>1.329</td>
</tr>
<tr>
<td>8. The course grading procedures were clearly explained.</td>
<td>4.965</td>
<td>1.428</td>
</tr>
<tr>
<td>9. Examinations (or equivalent) covered material studied in the course.</td>
<td>5.256</td>
<td>1.247</td>
</tr>
<tr>
<td>10. Course assignment &amp; examination directions were understandable.</td>
<td>5.234</td>
<td>1.077</td>
</tr>
<tr>
<td>11. Examination (or equivalent) feedback was timely &amp; adequate.</td>
<td>5.203</td>
<td>1.218</td>
</tr>
<tr>
<td>12. Assignments (papers, cases, problems, etc.) were related to course content.</td>
<td>5.332</td>
<td>0.0970</td>
</tr>
<tr>
<td>13. Assignment (papers, cases, problems, etc.) feedback was timely &amp; adequate.</td>
<td>5.232</td>
<td>1.144</td>
</tr>
<tr>
<td>14. Student questions were clearly &amp; adequately answered.</td>
<td>5.148</td>
<td>1.316</td>
</tr>
<tr>
<td>15. A productive learning environment was maintained for each session.</td>
<td>5.162</td>
<td>1.302</td>
</tr>
<tr>
<td>16. The instructor was reasonably available for consultation.</td>
<td>5.203</td>
<td>1.205</td>
</tr>
<tr>
<td>17. The instructor appeared to be well prepared for each session.</td>
<td>5.374</td>
<td>0.984</td>
</tr>
<tr>
<td>18. The instructor spoke clearly enough to be understood.</td>
<td>5.302</td>
<td>1.043</td>
</tr>
<tr>
<td>19. Examinations (or equivalent) were used to help learning occur.</td>
<td>5.148</td>
<td>1.410</td>
</tr>
<tr>
<td>20. Assignments (papers, cases, problems, etc.) helped learning occur.</td>
<td>5.145</td>
<td>1.377</td>
</tr>
<tr>
<td>21. The mix of teaching methods used helped learning occur.</td>
<td>4.978</td>
<td>1.675</td>
</tr>
<tr>
<td>22. The textbook(s) and/or handouts helped learning occur.</td>
<td>4.916</td>
<td>1.944</td>
</tr>
</tbody>
</table>

Note: For all items, a six point Likert style response option set was employed. For items 1-8, response options ranged between "very strongly disagree" (1) to "very strongly agree" (6). For items 4-22, response options ranged from "never" (1) to "always" (6).
Retained items loaded highly on only one factor. Four factors were identified: feedback, delivery of instruction, academic administration, and assessment of student learning. Cashin's model posited three of the four identified factors (i.e., dimensions).

The Factors

Factor loadings, sorted by factor, are presented in Table 3. It was unexpected to find the "feedback" factor which was originally envisioned to comprise the delivery of instruction dimension. The three items (11 to 13) loading on the feedback factor were related to the tendency of instructors to give clear feedback to students concerning their performance on various aspects of the course. The "delivery of instruction" factor was defined by items (14 to 18) which could be construed as relating to instructor rapport. Item 16 loaded on the delivery of instruction factor.

Items 4 to 8 loaded as expected and thus defined the "academic administration" factor. While
it can be accurately argued that Cashin's academic administration dimension can include other attributes, students seem most competent to differentiate selected course characteristics which are under the instructor's control. Three items (19 to 21) comprised the "assessment of student learning" factor. It does appear that assignments, examinations, and mix of teaching methods did engender student learning.

The factor correlation matrix is found in Figure 1. The factor correlations are rather high which poses a problem for the present index. Such high correlations suggest the presence of a single

<table>
<thead>
<tr>
<th>Table 3: Item Communality Estimates and Factor Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor</strong></td>
</tr>
<tr>
<td>Factor 1: Feedback</td>
</tr>
<tr>
<td>13. Assignment feedback was timely and adequate.</td>
</tr>
<tr>
<td>11. Examination (or equivalent) feedback was timely and adequate.</td>
</tr>
<tr>
<td>12. Assignment (papers, cases, problems, etc.) were related to course content.</td>
</tr>
<tr>
<td>Factor 2: Delivery of Instruction</td>
</tr>
<tr>
<td>17. The instructor appeared to be well prepared for each session.</td>
</tr>
<tr>
<td>18. The instructor spoke clearly enough to be understood.</td>
</tr>
<tr>
<td>15. A productive learning environment was maintained for each session.</td>
</tr>
<tr>
<td>16. The instructor was reasonably available for consultation.</td>
</tr>
<tr>
<td>14. Student questions were clearly and adequately answered.</td>
</tr>
<tr>
<td>Factor 3: Academic Administration</td>
</tr>
<tr>
<td>4. Course learning objectives were clearly stated and explained.</td>
</tr>
<tr>
<td>7. The syllabus clearly explained course organization and expectations</td>
</tr>
<tr>
<td>5. Course content was related to learning objectives.</td>
</tr>
<tr>
<td>6. Course organization was logical and understandable.</td>
</tr>
<tr>
<td>8. The course grading procedures were clearly explained.</td>
</tr>
<tr>
<td>Factor 4: Assessment of Student Learning</td>
</tr>
<tr>
<td>20. Assignment (papers, cases, problems, etc.) helped learning occur.</td>
</tr>
<tr>
<td>19. Examinations (or equivalent) were used to help learning occur.</td>
</tr>
<tr>
<td>21. The mix of teaching methods helped learning occur.</td>
</tr>
</tbody>
</table>
dimension which is highly unlikely. It is more logical to suspect that a strong halo effect is operating.

\[
\begin{array}{cccc}
\text{Factor} & 1 & 2 & 3 & 4 \\
\text{Factor 1} & - & & & \\
\text{Factor 2} & .88 & - & & \\
\text{Factor 3} & .72 & .75 & - & \\
\text{Factor 4} & .88 & .94 & .76 & - \\
\end{array}
\]

Figure 1: Factor Correlation Matrix

Predicting Global Satisfaction

A test of the usefulness of such an index is the extent to which the various identified dimensions could predict global satisfaction with the course and the instructor. Presented in Figure 2 are correlations between the identified dimensions and global items. All correlations are significant at alpha = .01. Such correlations were expected.

It was originally intended that stepwise regression analysis using the identified dimensions to predict the global items would be conducted to investigate further the nature of variable relationships. However, such was precluded because of multicollinearity due to the high correlations among the predictor variables.

\[
\begin{array}{ccccccc}
\text{Item 1} & \text{Item 2} & \text{Item 3} & \text{Feedback} & \text{Delivery} & \text{Admin} & \text{Assessment} \\
\text{Item 1} & 1.00 & & & & & \\
\text{Item 2} & .78 & 1.00 & & & & \\
\text{Item 3} & .60 & .73 & 1.00 & & & \\
\text{Feedback} & .55 & .63 & .57 & 1.00 & & \\
\text{Delivery} & .60 & .70 & .67 & .84 & 1.00 & \\
\text{Admin} & .74 & .83 & .74 & .68 & .71 & 1.00 \\
\text{Assessment} & .60 & .69 & .59 & .79 & .86 & .69 & 1.00 \\
\end{array}
\]

Figure 2: Correlations Between Model Dimensions and Global Items (all correlations, p < .01).

Discussion

Evidence for the existence of dimensions similar to those postulated was found, as well as evidence that the dimensions were related to the global rating items. Once revised, it should be
possible to use this or a similar instrument to assess an instructor’s teaching effectiveness and provide information for improving teaching. Cashin (1990a, p. 113-121) did report that students in differing academic fields rated instructors differently. However, Marsh and Hocevar (1991) reported that the factor structure of the SEEQ remained constant across 21 different academic fields. Such conflicting findings reinforce the Abrami, d’Apollonia, and Cohen (1990) recommendation that reliability, validity, and utility be assessed at the local level.

The high factor correlations presents a problem for the present index. It is most unlikely that teaching is a unidimensional construct. Thus, an alternative explanation is that a strong halo effect is operating and student differentiation among the dimensions is obscured. It is possible that in the present case (as in other institutions) that the SRTEs were completed in haste by untrained raters. Steps that can reduce the suspected halo effect include (a) administering the SRTEs under more controlled circumstances, (b) rewriting items to a simpler format, or (c) training raters.

It is also possible that the envisioned dimensions were inadequately defined; thus student ability to differentiate between dimensions was compressed. While Cashin’s theory is attractive, it is underdeveloped. There were no formal definitions proffered for any of the dimensions. However, he does offer suggestions as to the types of data which may be collected and which then could be used to help frame definitions. Feldman (1988) has identified 22 dimensions of teaching drawn from a meta-analysis of dozens of empirical reports on student ratings of teaching effectiveness. Presented in Appendix A are Feldman’s teaching dimensions. In order to develop definitions for each dimension relevant to Cashin’s student perspective, each of Feldman’s (1988) dimensions were labeled dimensional attributes and integrated, based on logical analysis, into Cashin’s student perspective dimensions (Appendix B). Thus, based on this integration, the following dimensional definitions, relevant to the student perspective, were drafted:

*Effective instructional delivery.* Hallmarks of effective instructional delivery include the stimulation of student interest in the subject fostered by an enthusiastic, well prepared instructor whose presentations are clear and understandable to students. Effective instructional delivery is provided by an instructor who is (a) an effective communicator with students; (b) aware of the learning level, generally, within his or her classroom; (c) establishes reasonably good rapport with students; (d) encourages students to take self-responsibility for
their own learning; (e) provides frequent feedback to students while attempting to answer questions fully and to involve students in class activities and discussions; and (f) characterized by a concerned and helpful attitude.

Assessment of student learning. This dimension is defined within the context of student learning. The assessment of student learning entails the examination of: (a) the usefulness of course instructional strategies in fostering student learning (e.g., assignments, tests, homework, readings, teaching methods, visual aids, etc.); (b) the general impact of instruction on students; (c) the quality of feedback to students to improve learning; (d) the perceived intellectual challenge of the course; and (e) whether or not students were held to high performance standards.

Administrative requirements (Academic Administration). Within the student perspective, effective academic administration entails a general understanding on the part of the student as to whether or not (a) the course organization was logical and understandable; (b) course learning objectives and student requirements were clear; (c) students grasp the relationship between the course and their broader education; (d) course content was related to course learning objectives; and (e) the instructional strategies (e.g., assignments, homework, readings, etc.) were related to course learning objectives and/or content.

No definition of the availability to student dimension was offered as earlier unpublished research, conducted by the authors into whether or not students could differentiate between the dimensions, as proposed by Cashin (1989), suggested that the availability dimension was in fact an attribute of the instructional delivery and academic administration dimensions. Additionally these factor analytical studies found that students could differentiate between three of Cashin's four dimensions (not availability), even with less substantial definitions than those offered above. In one of the unpublished studies, two items loaded on what could be labeled as an availability factor, but was determined to be unstable and was deleted. A revised index, based on the above definitions, is found in Appendix C. To further aid SRTE interpretation and use, Cashin (1990) has provided a set of guidelines for the use of SRTE data in faculty evaluation and development (Appendix D).

An empirically validated theoretical model built upon such a framework as advanced by Cashin can potentially improve teaching effectiveness assessment, faculty development efforts, and ultimately institutional academic effectiveness assessment. Once fully, developed such a theory can guide item pool development, improve administrative decision-making, and help faculty improve their teaching. However tentative, these data do suggest that such a model may be possible.
References


Office of Instructional Resources (no date). ICES Its Rationale and Description (Newsletter No. 2). Urbana-Champaign, IL: Office of Instructional Resources, University of Illinois.


Appendix A:
Feldman's Instructional Dimensions & Sample Items

Teachers stimulation of interest in the course and subject matter. (1)
Teachers enthusiasm for subject or teaching. (2)
Teachers knowledge of the subject. (3)
Teachers intellectual expansiveness and intelligence. (4)
Teachers preparation and organization of the course. (5)
Clarity and understandableness. (6)
Teachers elocutionary skills. (7)
Teachers sensitivity to and concern with class level and progress. (8)
Clarity of course objectives and requirements. (9)
Nature and value of the course material including its usefulness and relevance. (10)
Nature and usefulness of supplementary materials and teaching aides. (11)
Perceived outcome or impact of instruction. (12)
Instructors fairness; impartiality of evaluation of students; quality of examinations. (13)
Personality characteristics of the teacher. (14)
Nature, quality, and frequency of feedback from the teacher to the students. (15)
Teachers encouragement of questions and discussion and openness to opinion of others. (16)
Intellectual challenge and encouragement of independent thought by the teacher & course. (17)
Teachers concern and respect for student; friendliness of the teacher. (18)
Teacher availability and helpfulness. (19)
Teacher motivates students to do their best; high standards of performance required. (20)
Teachers encouragement of self-initiated learning. (21)
Teachers productivity in research and related activities. (22)
Appendix B: The Student Perspective Dimensional Attributes and Illustrative General Concept Items

**Delivery of Instruction: Dimensional Attributes**

Delivery: Stimulation of Student Interest (1)
Delivery: Teacher Enthusiasm (2)
Delivery: Teacher Knowledge (3) [Students unable to assess]
Delivery: Instructor Preparation (5)
Delivery: Presentation Clarity & Understandableness (6)
Delivery: Instructor's Elocutionary Skills (7)
Delivery: Learning Level & Process Awareness (8)
Delivery: Personal Characteristics (14)
Delivery: Frequency of Feedback (15)
Delivery: Class Discussions (16)
Delivery: Questions, Answers/Explanations (16)
Delivery: Teacher Concern for Students (18)
Delivery: Teacher Helpfulness (19)
Delivery: Self-initiated Learning Encouragement (21)

**Assessment of Student Learning: Dimensional Attributes**

Assessment: Usefulness of Course Instructional Materials (11)
   [Includes assignments, tests, homework, lab reports, & readings]
Assessment: Usefulness of Course Instructional Materials (11)
   [Includes teaching methods, visual aides, group work, etc.]
Assessment: Impact of Instruction (12)
Assessment: Quality of Assignment & Exam Items, etc. (13) [Students unable to assess]
Assessment: Grading Fairness (13) [Students unable to assess]
Assessment: Quality of Feedback (15)
Assessment: Intellectual Challenge (17)
Assessment: High Performance Standards (20)

**Academic Administration: Dimensional Attributes**

Administration: General Course Design (5)
Administration: Clarity of Course Objectives and Requirements (9)
Administration: Usefulness of Course Material (e.g., content) (10)
Administration: Nature of Course Material (e.g., content) (10)
Administration: Nature of Instructional Materials (11)
   [Includes assignments, tests, homework, lab reports, & readings]

**Supplemental**

Administration: Class Management
Administration: Lab Safety
Administration: Laboratory Equipment & Supplies
Appendix C: Student Assessment of Teaching Effectiveness
Student Assessment of Teaching Effectiveness
Department of Evening Classes
The University of Georgia

Directions: Please read each statement carefully and select one answer for each question, using the following scale:

- Strongly Disagree, circle 1.
- No Opinion/Neutral, circle 3.
- Strongly Agree, circle 5.
- Agree, circle 4.
- Disagree, circle 2.

First, please tell us what you generally think about the instructor and course.

1. OVERALL, the instructor was effective in helping me learn.
2. OVERALL, the course was effective in helping me learn.

Next, please tell use what you think about each of the following instructional or course characteristics.

3. The instructor appeared interested in teaching the course.
4. The instructor was usually well prepared for each class.
5. The instructor's presentations were clear and understandable.
6. The instructor is an effective communicator.
7. The instructor tried different approaches to explain concepts, content, skills, etc., when not understood.
8. The instructor established good rapport with students in the classroom.
9. Assignments and tests were reviewed and returned in a reasonable time.
10. Students were encouraged to participate in class discussions and activities.
11. The instructor tried to clearly and fully answer each question.
12. The instructor appeared interested in whether or not students learned.
13. The instructor was reasonably available to help students if requested.
14. Students were encouraged to take responsibility for their own learning.
15. The required content and/or skills were challenging to learn.
16. Students must perform well to earn a high grade.
17. The instructor explained how the course related to students' education.
18. The course organization was logical and understandable.
19. Course learning objectives, grading procedures, and student requirements were clearly stated and explained.
20. Course content appeared to be related to stated learning objectives.
21. Course readings, assignments, and tests were related to course content.
22. The instructor stimulated my interest in this subject.
23. The readings, assignments, and/or tests helped me learn the required information and/or skills.
24. I think I have achieved the course's learning goals or objectives.
25. The teaching methods used by the instructor helped me learn.
26. The feedback received on assignments and tests showed me where I needed to improve.

These last few questions ask you to describe yourself. Please answer each question.

27. What is your gender? 1 = Male, 2 = Female
28. What is your ethnic status? 1 = Native American, 2 = Asian or Pacific Islander, 3 = Black, 4 = Hispanic, 5 = White, 6 = Multiracial

Please continue on the back of this sheet.
30. What is your class rank? 1 = Freshman, 2 = Sophomore, 3 = Junior, 4 = Senior, 5 = Graduate, 6 = Irregular, 7 = Transient.
31. What is your current GPA, if established? 1 = <2.00, 2 = 2.01-2.49, 3 = 2.50-2.99, 4 = 3.00-3.49, 5 = 3.50-4.00.
32. Why are you taking this course? 1 = Required, 2 = Elective, 3 = Advisor suggested, 4 = Interesting subject, 5 = Instructor's reputation, 6 = Want to improve GPA.
33. How motivated were you to perform well in this course? 1 = Very Highly, 2 = Highly, 3 = Average, 4 = Poorly, 5 = Very Poorly.
34. What percentage of the class did you attend? 1 = 0-19, 2 = 20-39, 3 = 40-59, 4 = 60-79, 5 = 80-100.
35. How many hours per week did you study for this course? 1 = 0-2, 2 = 3-5, 3 = 6-8, 4 = 9-11, 5 = 12+

Use the space below to make comments. Please make specific recommendations to improve the course or to improve teaching effectiveness.

Instructor's Name: ____________________________________________________________
Course Prefix & Number: __________________________ Quarter: __________ Date: __________

Thank you for your time and effort.
Appendix D:
Recommendations for Using Student Ratings of Teaching

General Considerations

1. Use multiple sources of data about a faculty member's teaching if you are serious about accurately evaluating or improving teaching.
2. Do use student rating data as one source of data about effective teaching.
3. Discuss and decide upon the purpose(s) that the student rating data will be used for before any student rating form is chosen or any data are collected.

The System

4. To obtain reliable student rating data collect data from at least ten raters if this is possible.
5. To obtain representative student rating data from at least two-thirds of the class.
6. To generalize from student rating data to an instructor's overall teaching effectiveness, sample across both courses and across time.
7. For improvement, develop a student rating system that is flexible.
8. Provide comparative data, preferably for all the items. Student ratings tend to be inflated.
9. Discuss and decide what controls for bias will be included in your system.
10. Do not give undue weight to: the instructor's age, sex, teaching experience, personality or research productivity; the student's age, sex, level (freshman, etc.), grade-point-average, or personality; or the class size or time of day when it was taught.
11. Take into consideration the students' motivation level when interpreting student rating data.
12. Decide how you will treat student ratings from different course levels, e.g., freshman, graduate, etc.
13. Decide how you will treat student ratings from different academic fields.
14. For improvement, develop a system that is diagnostic.
15. Develop a system that is interpretable.

The Form

16. For evaluation, use a few global item or summary items or scores.
17. Use the short evaluation form (or items) in every class every term.
18. Use a long, diagnostic form in only one course per term—in the course that the instructor wishes to focus upon for improvement.
19. For improvement, use items that require as little inference as possible on the part of the student rater and as little interpretation as possible on the part of the instructor.
20. For improvement, do not use a single standard set of items for every class. Provide a pool of items or some kind of weighting system.
21. Use a 5-point to 7-point scale.
22. In the analysis of the results, report computations only to the first decimal place.
23. Do not overinterpret the data, allow for a margin of error.

24. Use frequency distributions—what number or percent of the students rated item “1” or “2,” etc. These are more understandable to most faculty.

25. For improvement, ask for open-ended as well as quantitative ratings.

26. Use open-ended comments only for improvement.

Administration

27. For evaluation, develop standardized procedures covering all relevant aspects of your student rating system and monitor that the procedures are followed.

28. For evaluation, administer the ratings about the second week to the last week of the term.

29. Develop standardized instructions that include the purpose(s) for which the data will be used and who will receive what information, and when.

30. Instruct the students not to sign their ratings.

31. The instructor may hand out the rating forms and read the standardized instructions, but the instructor should leave the room until the students have completed the ratings and they are collected.

32. The ratings should be collected by a neutral party and the data taken to a predetermined location—often to where they are scored—and they should not be available to the instructor until the grades are turned in.

Interpretation

33. Develop a written explanation of how the analyses of the student ratings are to be interpreted.

34. Appoint a faculty member to serve as instructional consultant to help faculty interpret their results and to improve teaching.

I. DOCUMENT IDENTIFICATION:

Title: Assessing Teaching Effectiveness in a Liberal Arts College: The Student Perspective,

Author(s): Charles D. Hale, Charlene Herreid & Gordon Waugh

Publication Date: May 1996

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following two options and sign at the bottom of the page.

Level 1

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

____Sample____

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 2

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

____Sample____

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Printed Name/Position/Title: Charles Hale

Telephone: 352-588-8309
FAX: 352-588-8161
E-Mail Address: cdha106
Date: 9-23-97
III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:

Address:

Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:

Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC CLEARINGHOUSE ON HIGHER EDUCATION
THE GEORGE WASHINGTON UNIVERSITY
ONE DUPONT CIRCLE, SUITE 630
WASHINGTON, D.C. 20036-1183

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2d Floor
Laurel, Maryland  20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

(Rev. 6/96)