DOCUMENT RESUME

ED 076 166   HE 004 181

AUTHOR   Aleamoni, Lawrence M.
TITLE   Evaluation by Students to Identify General
                Instructional Problems.
INSTITUTION   Illinois Univ., Urbana. Office of Instructional
                Resources.
PUB DATE   [73]
NOTE   9p.

EDRS PRICE   MF-$0.65 HC-$3.29
DESCRIPTORS   *Course Evaluation; *Effective Teaching; *Evaluation
                Methods; Faculty Evaluation; *Higher Education;
                *Teacher Evaluation; Teaching Quality
IDENTIFIERS   *Illinois Course Evaluation Questionnaire

ABSTRACT
        Of the various systems developed for student
evaluation of course and instructor, the Illinois Course Evaluation
Questionnaire (CEQ) has perhaps the most extensive reliability and
validity data to support it as well as the most extensive norm data
base. The CEQ is used to collect student attitudes towards a course
and instructor and its purpose is to enable faculty members to
collect evaluative information about their teaching. The CEQ can also
be used to provide feedback to administrators if it is couched in a
total instructional evaluation scheme that may consist of peer
evaluation, supervisor evaluation, classroom visitation, course
material evaluation, future student success and achievement data.
(Author/HS)

Evaluation by Students to Identify General Instructional Problems[1]

Lawrence M. Aleamoni, Head
Measurement and Research Division
Office of Instructional Resources
307 Engineering Hall
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

In the past few years as a result of the 1970 student strikes and the emphasis on accountability, course and instructor evaluation has been placed in the spotlight. In an attempt to build a total instructional evaluation system, a great deal of emphasis has been placed on student evaluations of course and instructor. In order for student evaluations to be considered an integral part of a total instructional evaluation system, they must be both reliable and valid.

Of the various systems developed for student evaluation of course and instructor, the Illinois Course Evaluation Questionnaire (CEQ) has perhaps the most extensive reliability and validity data to support it as well as the most extensive norm data base. Norm data have been collected continuously since 1966 at the University of Illinois, Urbana-Champaign campus. The CEQ is used to collect student attitudes towards a course and instructor and its purpose is to enable faculty members to collect evaluative information about their teaching. Once the instructor has used the CEQ and submitted the forms for analysis, two copies of the results are returned only to the instructor. However, pooled results for entire courses consisting of many sections may also be obtained with instructor identification omitted. As

---

[1] Part of a symposium presentation entitled "A Plan for the Comprehensive Evaluation of College Teaching" at the American Educational Research Association Convention in New Orleans, Louisiana, February 27, 1973.

the number of measures on each course is increased, it becomes possible to obtain a relatively stable indication of the difference between courses. This aids in the interpretation of the actual differences between an obtained section score for a particular instructor and the average scores for all the sections represented in that course.

The analysis of item inter-relationships and the subscore inter-relationships indicated that no one element, related to a course, disproportionately influenced the students' evaluation of the course (Spencer & Aleamoni, 1969). It appears that there is a "general course attitude" cultivated by the student as he is exposed to previous student's comments, the instructor, the textbook, the course, etc., and this is the framework from which he responds when answering the CEQ items.

It would seem, on the basis of three validity studies (Stallings & Spencer, 1967; Swanson & Sisson, 1971; Aleamoni & Yimer, 1972), the face validity of the CEQ, and its high reliability, that extremely low scores on a particular subscore should indicate problem areas in an instructor's teaching procedure. Whereas, stable high scores should point to an effective instructional program as viewed by students. All available validating evidence (both published and unpublished studies) to date, indicates that the CEQ does indeed identify courses that are considered to be excellent or poor.

After using the CEQ, the instructor receives results (see Appendix A) which allow him to compare his course item means to institutional course item means (via deciles) and his course subscale means to norm subscale means categorized by (a) rank of instructor, (b) level of course, (c) institution, (d) college, and (e) all institutions that have used the CEQ throughout the United States. The subscale results allow the instructor to obtain an indication of major areas of strengths and weaknesses in the course. Once

the areas of weakness have been identified by the subscales, then looking
at the item results helps to focus on the more specific problem areas. The
CEQ items are not completely diagnostic but do serve to elicit diagnostic
responses from the instructor teaching the course. It provides a means
whereby some evaluation of the teaching process can occur; other means can
be arranged and are available such as asking more diagnostic questions in
the optional item section available on the CEQ form, or having peers sit in
on actual class sessions, etc. It is important to recognize, however, that
student opinions are in existence and do affect learning--and they do provide
a source of quite reliable and valid data relative to the effectiveness of
instruction (Costin, Greenough & Menges, 1971).

In order to provide instructors with items that may be more relevant
or diagnostic for their particular courses, a catalog of items was generated
by the Measurement and Research Division of the Office of Instructional
Resources at the University of Illinois, Urbana-Champaign campus. The items
were gathered from all existing sources such as institutional, national, de-
partmental, and individual instructor questionnaires. They were then restated
so that the response categories of strongly agree (SA), agree (A), disagree (D),
and strongly disagree (SD) would apply. This then made it possible for those
items to be used in the "Optional Item" section of the CEQ (see Appendix B).

This collection of some 270 items was divided into 19 categories consisting
of: (a) instructor contribution, (b) attitude toward students, (c) student
outcomes, (d) relevance of course, (e) use of class time, (f) organization and
presentation, (g) clarity of presentation, (h) instructor characteristics,
(i) interest of presentation, (j) expectations and objectives, (k) behavioral
indications of course attitude, (l) general attitude toward instructor, (m)
speed and depth of coverage, (n) out-of-class, (o) examinations, (p) visual
aids, (q) grading, (r) assignments, and (s) laboratory and recitation.

The response to the availability of the catalog of optional items was gratifying in that it was not finished until December 12, 1972, less than four weeks before the end of the fall semester. Of 1414 course sections using the CEQ during fall semester 1972, approximately 313 made use of the optional item section.

After the instructor has decided to use the CEQ and/or any optional items of his choice, it is then up to him to decide what to do with the data. If he feels that the interpretation manual (Aleamoni, 1972) and abbreviated interpretation sheets are not sufficient to help him identify areas that may need improvement in the course, he can then arrange for a conference with one of the members of the Measurement and Research Division staff. Such a conference would begin with a close scrutiny of the CEQ sub-scale results to see if any problems existed based on the norm data available. If a problem area was identified (such as Method of Instruction) then a close look at the items making up that subscale would be in order. If, in the discussion with the instructor the source of difficulty is identified, then the discussion would shift to possible ways of trying to resolve the difficulty. If, on the other hand, the source of difficulty cannot be identified using the existing items and the instructor's recall, then pro-cedures (such as the use of optional items that are much more diagnostic) would be explored to be able to identify the specific problem.

It has been through a process such as this that instructors have been able to use student evaluations to identify instructional problems and then rectify them. Obviously, the success or failure of such a venture rests solely with the instructor and his willingness to both gather and use the data provided him.

The CEQ can also be used to provide feedback to administrators (deans, department heads, etc.) if it is couched in a total instructional evaluation scheme which may consist of peer evaluation, supervisor evaluation, classroom visitation, course material evaluation, future student success and achievement data, etc. Student course and instructor evaluation data should not be used alone in evaluating instructional effectiveness in rank, pay, and tenure decisions as it is not completely diagnostic of all elements in the teaching or instruction domain. The actual weight applied to student evaluations by administrators varies from 0 to 100% depending upon the department, college and institution. What weight student evaluations should ideally carry is contingent upon the number and quality of other elements used in the instructional evaluation scheme.

References

Aleamoni, L. M. Results and Interpretation Manual: Illinois Course Evaluation Questionnaire, Research Report No. 331 . Urbana, Illinois: Measurement and Research Division, Office of Instructional Resources, University of Illinois at Urbana-Champaign, 1972.

Aleamoni, L. M. & Yimer, M. An investigation of the relationship between colleague rating, student rating, research productivity, and academic rank in rating instructional effectiveness. Research Report No. 338. Urbana, Illinois: Measurement and Research Division, Office of Instructional Resources, University of Illinois at Urbana-Champaign, 1972.

Costin, F., Greenough, W. T. & Menges, R. J. Student ratings of college teaching: Reliability, validity, and usefulness. Review of Educational Research, 1971, 41(5), 511-535.

Spencer, R. E. & Aleamoni, L. M. The Illinois Course Evaluation Questionnaire: A description of its development and a report of some of its results. Research Report No. 292. Urbana, Illinois: Measurement and Research Division, Office of Instructional Resources, University of Illinois at Urbana-Champaign, 1969.

Stallings, W. M. & Spencer, R. E. Ratings of instructors in Accountancy 101 from video-tape clips. Research Report No. 265. Urbana, Illinois: Measurement and Research Division, Office of Instructional Resources, University of Illinois at Urbana-Champaign, 1967.

Swanson, R. A. & Sisson, D. J. The development, evaluation, and utilization of a departmental faculty appraisal system. Journal of Industrial Teacher Education, 1971, 9(1), 64-79.

Appendix A

RESULTS FOR THE OBJECTIVE ITEMS OF THE ADVISOR QUESTIONNAIRE

20140 ALEAMONI LARRE EDPSY 490 SECTION 4   ENROL=0005 FALL 1971   03620J

SEX
FEMALE MALE OMIT
0.20   0.20   0.60

MAJOR-MINOR
MAJOR MINOR OTHER OMIT
0.40   0.20   0.40   0.00

COURSE OPTION
REQ ELECT OMIT
0.40   0.40   0.20

PASS-FAIL
YES   NO   OMIT
0.00 0.60 0.40

STATUS
FRESH SOPH   JR   SP   GRAD OTHER OMIT
0.00   0.00 0.00 0.00 1.00   0.00   0.00

EXPECTED GRADE
A   B   C   D   E   OMIT
0.60   0.40 0.00 0.00 0.00 0.00

COURSE GRADE
A   B   C   D   E   OMIT
0.80   0.20 0.00 0.00 0.00 0.00

INSTRUCTOR GRADE
A   B   C   D   E   OMIT
1.00   0.00 0.00 0.00 0.00 0.00

ITEM	SA	A	D	SD	OMIT	BEST	MEAN	S.D.	DECL	0123456789
1.	0.00	0.00	0.40	0.60	0.00	SD	3.60	0.55	9	*
2.	0.00	0.00	0.40	0.60	0.00	SD	3.60	0.55	7	*
3.	0.60	0.40	0.00	0.00	0.00	SA	3.60	0.55	9	*
4.	0.40	0.60	0.00	0.00	0.00	SA	3.40	0.55	9	*
5.	0.80	0.20	0.00	0.00	0.00	SA	3.80	0.45	8	*
6.	0.60	0.40	0.00	0.00	0.00	SA	3.60	0.55	9	*
7.	0.80	0.20	0.00	0.00	0.00	SA	3.80	0.45	9	*
8.	0.00	0.20	0.40	0.40	0.00	SD	3.20	0.84	8	*
9.	0.20	0.80	0.00	0.00	0.00	SA	3.20	0.45	8	*
10.	0.00	0.00	0.40	0.60	0.00	SD	3.60	0.55	9	*
11.	0.00	0.00	0.20	0.80	0.00	SD	3.80	0.45	9	*
12.	0.20	0.80	0.00	0.00	0.00	SA	3.20	0.45	7	*
13.	0.60	0.40	0.00	0.00	0.00	SA	3.60	0.55	9	*
14.	0.00	0.00	0.40	0.60	0.00	SD	3.60	0.55	9	*
15.	0.00	0.00	0.20	0.80	0.00	SD	3.80	0.45	9	*
16.	0.80	0.20	0.00	0.00	0.00	SA	3.80	0.45	9	*

MERMAC -- TEST ANALYSIS AND QUESTIONNAIRE PACKAGE

*	*	0.00	0.00	0.20	0.80	0.00	SD	3.80	0.45	9	* *
17.	0.00	0.00	0.00	0.20	0.80	0.00	SD	3.80	0.45	9	*
18.	1.00	1.00	0.00	0.00	0.00	0.00	SA	4.00	0.00	9	*
19.	0.20	0.80	0.80	0.00	0.00	0.00	SA	3.20	0.45	6	*
20.	0.60	0.40	0.40	0.00	0.00	0.00	SA	3.60	0.55	9	*
21.	0.40	0.40	0.20	0.20	0.00	0.00	SA	3.20	0.84	8	* *
22.	0.40	0.60	0.60	0.00	0.00	0.00	SA	3.40	0.55	9	*
23.	0.20	0.00	0.00	0.20	0.60	0.00	SD	3.20	1.30	6	*
24.	0.00	0.00	0.00	0.20	0.80	0.00	SD	3.80	0.45	9	
25.	0.40	0.60	0.00	0.00	0.00	0.00	SA	3.40	0.55	8	*
26.	0.00	0.00	1.00	0.00	0.00	0.00	SD	3.00	0.00	8	*
27.	0.60	0.40	0.00	0.00	0.00	0.00	SA	3.60	0.55	9	*
28.	0.00	0.20	0.60	0.00	0.00	0.00	SD	2.80	0.45	1	*
29.	0.00	0.00	0.00	1.00	0.00	0.00	SD	4.00	0.00	9	
30.	0.00	0.00	0.20	0.00	0.00	0.00	SA	2.80	0.45	5	*
31.	0.00	0.00	0.00	1.00	0.00	0.00	SD	4.00	0.00	9	*
32.	0.00	0.40	0.60	0.60	0.00	0.00	SD	2.50	0.55	1	*
33.	0.00	0.00	0.00	0.60	0.00	0.00	SD	3.60	0.55	9	*
34.	0.40	0.00	0.20	0.80	0.00	0.00	SD	3.80	0.45	9	*
35.	0.00	0.60	0.60	0.00	0.00	0.00	SA	3.40	0.55	8	*
36.	0.60	0.40	0.40	0.00	0.00	0.00	SA	3.60	0.55	9	* *
37.	0.00	0.00	0.40	0.60	0.00	0.00	SD	3.60	0.55	9	* *
38.	0.00	0.20	0.60	0.40	0.00	0.00	SD	3.40	0.55	8	*
39.	0.00	0.40	0.60	0.00	0.00	0.00	SD	3.60	0.55	6	* *
40.	0.40	0.60	0.00	0.00	0.00	0.00	SA	3.40	0.55	9	* *
41.	0.00	0.20	0.60	0.20	0.00	0.00	SD	3.00	0.71	6	*
42.	1.00	0.00	0.00	0.00	0.00	0.00	SA	4.00	0.00	9	*
43.	0.00	0.80	0.00	0.20	0.00	0.00	SD	3.20	0.45	8	*
44.	0.00	0.20	0.40	0.40	0.00	0.00	SD	3.70	0.84	6	
45.	0.00	0.00	0.60	0.40	0.00	0.00	SD	3.40	0.55	9	*
46.	0.40	0.00	0.20	0.00	0.80	0.00	SD	3.20	0.45	9	*
47.	0.80	0.20	0.00	0.00	0.00	0.00	SA	3.80	0.45	9	*
48.	0.00	0.00	0.40	0.60	0.00	0.00	SD	3.60	0.55	9	
49.	0.40	0.60	0.00	0.00	0.00	0.00	SA	3.40	0.55	8	*
50.	0.60	0.40	0.00	0.60	0.00	0.00	SA	3.60	0.55	9	*

--SUBSCORE--	ITEMS	RESP	MEAN	S.D.	REL	RANK	LEVEL	INSTL	COLL	OVER-ALL
GENERAL ATTITUDE	8	1.00	3.65	0.48	0.90	NONE	9	9	NONE	9
METHOD	8	1.00	3.55	0.55	0.85	NONE	5	7	NONE	9
CONTENT	8	1.00	3.07	0.57	0.66	NONE	9	9	NONE	7
INTEREST	8	1.00	3.55	0.50	0.93	NONE	9	9	NONE	9
INSTRUCTOR	8	1.00	3.61	0.62	0.00	NONE	9	9	NONE	9
SPECIFIC ITEMS	10	1.00	3.40	0.64	0.09	NONE	9	9	NONE	9
TOTAL	50	1.00	3.48	0.60	0.93	NONE	9	9	NONE	9

SAMPLE SIZE = 5