

taining the need, content, and clientele for staff development programs and for the generation of instructional materials for the improvement of college teaching.

In addition, the rationale and methodology of the study can have important implications for two groups of investigators: those seeking to determine the relationships between additional teaching behaviors and positive college student learning outcomes and those seeking justification for the inclusion of any given teacher competency in a competency-based teacher education program.

Bibliography

Centra, J. A. 1972. *The utility of student ratings for instructional improvement*. PR 72-16. Educational Testing Service, Princeton.

Fahey, G. L. 1967. *A measure of college instructor behavior*. University of Pittsburgh.

Gage, N. L., Belgard, M., and Rosenshine, B. 1968. The teacher's effectiveness in explaining: *Evidence on its generality and correlation with pupils' ratings and attention scores*. Technical Rept. No. 4. Stanford Center for Research and Development in Teaching. Stanford University.

Holland, J. L., and Lutz, S. W. 1968. *The predictive validity of a student's choice of vocation*. Personnel and Guidance J. 46:428-434.

Hoyt, D. P. 1969. *Identifying effective teaching behaviors*. Kansas State University, Manhattan, Kansas. 136 pp.

Hoyt, D. P., Owens, R. E., and Grouling, T. 1973. *Interpreting student feedback on institution and courses*. Office of Educational Resources, Kansas State University, Manhattan, Kansas.

Johnson, J. A. 1967. *Instruction: From the consumer's view*. In Improving college teaching. Amer. Coun. on Ed., Washington, D.C.

Keefer, K. W. 1965. *Self-prediction of academic achievement by college students*. Thesis for degree of Ph.D., University of Tennessee. Unpublished.

McKeachie, W. J., and Lin, Y. 1973. *Student characteristics related to achievement in introductory psychology courses*. Brit. J. Ed. Psych. 43:70-71.

Rosenshine, B. 1971. *Teaching behaviors and student achievement*. National Foundation for Educational Research in England and Wales, Book Publishing Division, 2 Jennings Buildings, Thomas Avenue, Windsor, Berks SL41Q5, Great Britain.

Solomon, D., Bezdek, W. E., and Rosenberg, L. 1963. *Teaching styles and learning*. Center for the Study of Liberal Education for Adults, Chicago.

Solomon, D., Rosenberg, L., and Bezdek, W. E. 1964. *Teaching behaviors and student learning*. J. Ed. Psych. 55:23-30.

Walsh, W. B. 1967. *Validity of self-report*. J. Coun. Psych. 14:18-23.

Evaluation

Typical Faculty Concerns About Student Evaluation of Instruction

Abstract

Eight typical faculty concerns about the appropriateness of using student ratings of instructor and instruction are presented. Discussions of the answers to these concerns are presented using a plethora of research that spans at least 50 years. Finally, faculty members are asked to consider the eight concerns from the point of view of faculty evaluating students.

Lawrence M. Aleamoni

In the past few years there have been many proposals for evaluating instruction, and a few of them were also concerned with trying to relate evaluation to the improvement of instruction. Most proposals suggested the use of similar elements in the evaluation procedure. These include (a) judgment by student, peer, self, and supervisor (department head), and (b) judgments of course material, course content, course objectives, and quality of student learning. If, however, one looks for actual working models of instructional evaluation, it is immediately apparent that schemes involving systematic ratings by peer, supervisor, or self, or of material, content, etc., are rarely actualized. More often than not, the

student ratings of instructor and instruction appear as the only elements in any of the "working models," and there are many reasons one could cite for this. This paper, however, will focus specifically on eight typical faculty concerns about the appropriateness of using ratings of instructor and instruction. These are summarized below in terms of common observations frequently expressed by faculty.

Typical Faculty Concerns

1. Students cannot make consistent judgments concerning the instructor and instruction because of their immaturity, lack of experience, and capriciousness.
2. Only colleagues with excellent publication records and experience are qualified to evaluate their peer's instruction.
3. Most student rating schemes are nothing more than a popularity contest with the warm, friendly, humorous, easy-grading instructor emerging as the winner.
4. Students are not able to make accurate judgments until they have been away from the course and possibly away from the university for several years.
5. The student rating forms are both unreliable and invalid.

Aleamoni is director of the Office of Instructional Research and Development and professor of Educational Psychology, The University of Arizona, Tucson.

6. There are extraneous variables or conditions that can affect student ratings. Some of the more common ones are (a) the size of the class, (b) the sex of the student, (c) the time of day the course was offered, (d) whether the student was taking the course as a requirement or on an elective basis, (e) whether the student was a major or a nonmajor, (f) the term (or semester) the course was offered, (g) the level of the course (freshman, sophomore, junior, senior, graduate), and (h) the rank of the instructor (instructor, assistant professor, associate professor, full professor).

7. The grades or marks students receive in the course are highly correlated with their ratings of the course and the instructor.

8. Finally, a question that is frequently raised is "How can student evaluations possibly be used to improve instruction?"

Surprising as it may be, answers to these problems and questions can be found in a plethora of research that spans at least 50 years. Most of this research has been conducted using student evaluation (rating) questionnaires similar to the one presently used at the University of Illinois at Urbana-Champaign (UIUC) called the Illinois Course Evaluation Questionnaire (CEQ) (see Appendix A for a copy of the CEQ).

Before beginning to cite and summarize the research addressing each of the problems and questions above, I will present a brief description of the CEQ in order to provide a meaningful frame of reference for the various studies.

The CEQ

The CEQ is an instrument used to collect student attitudes and opinions toward a course (Aleamoni & Spencer, 1973; Aleamoni, Note 1). Its purpose is to enable faculty members to collect evaluative information about their teaching. The data are collected and processed by course section but may also be processed by course, department, college. Extensive normative data have been gathered on the CEQ over the past 10 years to provide the instructor with valuable comparisons on the instructional dimensions of (a) general course attitude, (b) course content, (c) method of instruction, (d) interest and attention of the students, and (e) the instructor. These instructional dimensions represent the subscales of the CEQ. The 23 items that make up the subscales provide some diagnostic feedback to the instructor through appropriate norm comparisons. For example, each instructor's results are compared with those of other instructors, of his/her own academic rank; teaching at the same course level; and in his/her own department, college, or university. In addition, comparisons are made to all instructors who have used the CEQ throughout the United States.

Student responses to the CEQ are anonymous, and two copies of the questionnaire results (see Appendix B for sample output) along with interpretative information are returned only to the instructor. Instructors may de-

cide to submit one copy to their department chairperson for rank, pay, and tenure consideration; but the CEQ is primarily used to provide feedback to instructors to indicate where potential problems may exist in the classroom. There is ample space on the CEQ form for instructors to utilize more specifically diagnostic items in identified problem areas. Conferences relative to the interpretation and utilization of results may be arranged at the instructor's convenience with consultants from the Measurement and Research Division of the Office of Instructional Resources at UIUC.

With that brief description of the CEQ as background, I will respond to the eight concerns.

1. Students cannot consistently judge instructor and instruction. There is ample evidence on this point dating back to 1924, according to Guthrie (1954), in which reliabilities of student ratings remain in the .80 to .90 range. More recent literature on the subject by Costin, Greenough, and Menges (1971) and Aleamoni (Note 2) has shown that well-developed instruments and administration procedures can consistently yield high reliabilities (i.e., reliabilities in the .90s).

2. Only colleagues with excellent publication records and experiences are qualified to evaluate their peer's instruction. Wherever a discussion about or a proposal for student ratings emerges, this statement can usually be heard. Recently, a well-known statistician presented just such an argument in *The American Statistician* (Demings, 1972). Fortunately (or unfortunately for those who believe such a contention) about the time that Professor Deming's article appeared, a study had been completed addressing that very topic. Aleamoni and Yimer (1973) found that colleague and student ratings were not significantly related to the instructor's research productivity: the correlations were .07 and -.04, respectively. In addition, there was ample evidence previously (Guthrie, 1954; Swanson & Sisson, 1971; Stallings & Spencer, Note 3) to show that colleague and student ratings were very highly correlated (from .63 to .70). The response to Professor Deming's article can be found in a later issue of *The American Statistician* (Aleamoni, 1972).

3. Most student rating schemes are nothing more than a popularity contest. Answers to this problem are presented from published and unpublished studies on the CEQ and *The Advisor* (a student-sponsored form) (Feldman, Note 4) at UIUC. The studies conducted in developing and utilizing the CEQ subscales (Aleamoni & Spencer, 1973) indicated that no single subscale (i.e., Method of Instruction) completely overlapped the other subscales. Basically, this means that an instructor who received a high decile rating on the Instructor subscale (made up of items like "The instructor seemed to be interested in students as persons.") would not be guaranteed high decile ratings on the other four subscales (General Course Attitude, Method of Instruction, Course

FORM 73 ILLINOIS COURSE EVALUATION QUESTIONNAIRE CEQ

Measurement and Research Division, Office of Instructional Resources The Board of Trustees of The University of Illinois, 1972

MARK:

SA IF YOU STRONGLY AGREE WITH THE ITEM

A IF YOU AGREE MODERATELY WITH THE ITEM

D IF YOU DISAGREE MODERATELY WITH THE ITEM

SD IF YOU STRONGLY DISAGREE WITH THE ITEM

1. MARK ONLY ONE RESPONSE PER ITEM USING PENCIL ONLY. (SA A D SD)

2. ERASE CHANGED ANSWERS CLEANLY AND COMPLETELY. (b) SA A D SD

INSTRUCTIONS:

1. WRITE THE COURSE CODE IN THE SPACE PROVIDED.

2. WRITE THE INSTRUCTOR'S NAME IN THE SPACE PROVIDED.

3. RATE THE FOLLOWING INSTRUCTOR BEHAVIOR USING THE FOLLOWING SCALE:

EXCELLENCE VERY GOOD GOOD FAIR POOR

SCALES:

GENERAL ATTITUDE: 1-5 (1=EXCELLENCE, 5=POOR)

METHOD: 1-5 (1=EXCELLENCE, 5=POOR)

CONTENT: 1-5 (1=EXCELLENCE, 5=POOR)

INTEREST: 1-5 (1=EXCELLENCE, 5=POOR)

INSTR. GENERAL: 1-5 (1=EXCELLENCE, 5=POOR)

INSTR. SPECIFIC: 1-5 (1=EXCELLENCE, 5=POOR)

TOTAL: 1-5 (1=EXCELLENCE, 5=POOR)

ITEM	MARKS				OMIT	BEST	MEAN	S.D.	DECL.	0123456789
	SA	A	D	SD						
1	SA A D SD	I would take another course that was taught this way.	24	47	48	25	26	27	28	29
2	SA A D SD	The instructor seemed to be interested in students as persons.	24	48	49	25	26	27	28	29
3	SA A D SD	I would have preferred another method of teaching in this course.	26	60	61	27	28	29	30	31
4	SA A D SD	It was easy to remain attentive.	26	60	61	27	28	29	30	31
5	SA A D SD	The instructor did NOT synthesize, integrate or summarize effectively.	26	61	62	27	28	29	30	31
6	SA A D SD	NOT much was gained by taking this course.	26	61	62	27	28	29	30	31
7	SA A D SD	The instructor encouraged the development of new viewpoints and appreciations.	30	53	54	31	32	33	34	35
8	SA A D SD	I learn more when other teaching methods are used.	31	54	55	32	33	34	35	36
9	SA A D SD	The course material seemed worthwhile.	32	55	56	33	34	35	36	37
10	SA A D SD	The instructor was excellent.	33	56	57	34	35	36	37	38
11	SA A D SD	The instructor demonstrated a thorough knowledge of the subject matter.	34	57	58	35	36	37	38	39
12	SA A D SD	I would rather NOT take another course from this instructor.	35	58	59	36	37	38	39	40
13	SA A D SD	It was a very worthwhile course.	36	59	60	37	38	39	40	41
14	SA A D SD	Some things were NOT explained very well.	37	60	61	38	39	40	41	42
15	SA A D SD	The course material was too difficult.	38	61	62	39	40	41	42	43
16	SA A D SD	This was one of my poorest courses.	39	62	63	40	41	42	43	44
17	SA A D SD	The instructor seemed to consider teaching as a chore or routine activity.	40	63	64	41	42	43	44	45
18	SA A D SD	It was quite interesting.	41	64	65	42	43	44	45	46
19	SA A D SD	I think that the course was taught quite well.	42	65	66	43	44	45	46	47
20	SA A D SD	The course content was excellent.	43	66	67	44	45	46	47	48
21	SA A D SD	Some days I was NOT very interested in this course.	44	67	68	45	46	47	48	49
22	SA A D SD	It was quite boring.	45	68	69	46	47	48	49	50
23	A A D SD	Overall, the course was good.	46	69	70	47	48	49	50	51

OPTIONAL ITEM SECTION II

70 A B C D E 71 A B C D E 72 A B C D E 73 A B C D E 74 A B C D E 75 A B C D E 76 A B C D E 77 A B C D E

78 A B C D E 79 A B C D E 80 A B C D E 81 A B C D E 82 A B C D E 83 A B C D E 84 A B C D E 85 A B C D E

PLEASE FILL OUT THE OTHER SIDE

Table 1. Results for the objective items on the CEQ Form 73 Questionnaire.

STATUS	FRESH	SOPH	JR.	SR.	GRAD	OTHER	OMIT			
		0.87	0.13	0.00	0.00	0.00	0.00	0.00		
EXPECTED GRADE	A	B	C	D	E	OMIT				
	0.40	0.53	0.07	0.00	0.00	0.00				
PASS-FAIL	YES	NO	OMIT							
	0.00	0.87	0.13							
COURSE OPTION	REQ	ELECT	OMIT							
	0.40	0.47	0.13							
MAJOR-MINOR	MAJOR	MINOR	OTHER	OMIT						
	0.13	0.00	0.80	0.07						
SEX	FEMALE	MALE	OMIT							
	0.07	0.37	0.07							
CONTENT RATING	EXCEL	VGD	GOOD	FAIR	POOR	VPR	OMIT			
	0.27	0.47	0.20	0.00	0.00	0.00	0.07			
INSTR RATING	EXCEL	VGD	GOOD	FAIR	POOR	VPR	OMIT			
	0.80	0.07	0.07	0.00	0.00	0.00	0.07			
COURSE RATING	EXCEL	VGD	GOOD	FAIR	POOR	VPR	OMIT			
	0.27	0.60	0.07	0.00	0.00	0.00	0.07			
SEMESTER	FALL	SPRING	SUMMER	OMIT						
	0.00	0.80	0.00	0.20						
ITEM	SA	A	D	SD	OMIT	BEST	MEAN	S.D.	DECL.	0123456789
1.	0.67	0.33	0.00	0.00	0.00	SA	3.67	0.49	9	*
2.	0.93	0.07	0.00	0.00	0.00	SA	3.93	0.26	9	*
3.	0.00	0.27	0.27	0.47	0.00	SD	3.20	0.86	8	*
4.	0.67	0.33	0.00	0.00	0.00	SA	3.67	0.49	9	*
5.	0.00	0.00	0.33	0.67	0.00	SD	3.67	0.49	9	*
6.	0.00	0.00	0.33	0.67	0.00	SD	3.57	0.49	8	*
7.	0.33	0.53	0.13	0.00	0.00	SA	3.20	0.68	6	*
8.	0.00	0.27	0.47	0.27	0.00	SD	3.00	0.76	6	*
9.	0.60	0.40	0.00	0.00	0.00	SA	3.60	0.51	9	*
10.	0.93	0.07	0.00	0.00	0.00	SA	3.93	0.26	9	*
11.	0.87	0.13	0.00	0.00	0.00	SA	3.87	0.35	8	*
12.	0.00	0.00	0.20	0.80	0.00	SD	3.80	0.41	9	*
13.	0.47	0.53	0.00	0.00	0.00	SA	3.47	0.52	8	*
14.	0.00	0.13	0.40	0.47	0.00	SD	3.33	0.72	9	*
15.	0.00	0.00	0.53	0.47	0.00	SD	3.47	0.52	8	*
16.	0.00	0.07	0.33	0.60	0.00	SD	3.53	0.64	8	*
17.	0.00	0.00	0.27	0.73	0.00	SD	3.73	0.46	9	*
18.	0.53	0.47	0.00	0.00	0.00	SA	3.53	0.52	9	*
19.	0.53	0.47	0.00	0.00	0.00	SA	3.53	0.52	8	*
20.	0.33	0.60	0.07	0.00	0.00	SA	3.27	0.59	8	*
21.	0.00	0.20	0.40	0.40	0.00	SD	3.20	0.77	9	*
22.	0.00	0.00	0.20	0.80	0.00	SD	3.80	0.41	9	*
23.	0.73	0.27	0.00	0.00	0.00	SA	3.73	0.46	9	*

SUBSCORE	ITEMS	RESP	MEAN	S.D.	REL RANK	LEVEL INSTI	COLLEGE	OVER-ALL
GENERAL ATTITUDE	4	1.00	3.60	0.33	0.35	8	9	8
METHOD	4	1.00	3.35	0.53	0.93	8	9	8
CONTENT	4	1.00	3.42	0.45	0.86	9	9	9
INTEREST	4	1.00	3.33	0.40	0.77	9	9	9
INSTR. GENERAL	2	1.00	3.87	0.30	0.70	9	9	9
INSTR. SPECIFIC	5	1.00	3.68	0.27	0.49	9	9	9
TOTAL	23	1.00	3.56	0.30	0.90	9	9	9

SAMPLE SIZE = 15

Content, and Interest and Attention). To explore this problem more fully, I reviewed the written comments made by students on both the CEQ and *The Advisor* and then compared them to their objective responses. The results indicated that students would frankly praise instructors for their warm, friendly, humorous manner in the classroom; but if their courses were not well organized or their methods of stimulating students to learn were poor, the students would equally frankly criticize them in those areas. When these comments were compared to the objective measures in the same areas, a high degree of relationship was observed. This evidence, in addition to that presented by Costin et. al. (1971) and Grush and Costin (1975), indicates that students are discriminating judges and not easily fooled by the good "showman" who is lacking in the other instructional qualities.

4. Students cannot make accurate judgments until they are out of the course and away from the university for several years. This point is repeatedly raised by faculty and was recently presented by McKeachie (1969). It is very difficult to obtain a comparative and representative sample in longitudinal follow-up studies. The sampling problem is further compounded by the fact that almost all student attitudinal data relating to a course or instructor are gathered anonymously. Most studies in this area, therefore, have relied on surveys of alumni and/or graduating seniors.

Earlier studies by Drucker and Remmers (1950, 1951) showed that alumni who were out of school five to ten years rated instructors much the same as students currently enrolled. More recent evidence by Aleamoni and Yimer (Note 5) further substantiates the earlier findings. This evidence indicates, contrary to popular belief and speculation, that students are very perceptive in their judgments and are in substantial agreement with peers who have been out of the course and away from the university for several years. A very carefully controlled follow-up study, however, needs to be conducted to fully answer this problem.

5. The student rating forms are both unreliable and invalid. This problem can be divided into two portions, one concerning the reliability of student rating forms and the other the validity of student rating forms.

The research literature is replete with studies that answer the question of the reliability of student rating forms. Almost all of the instruments which have been carefully constructed and tested by professionals yield reliabilities at the level of .80 and .90 on the subscales as well as the total instrument (Costin et al., 1971). Reliabilities computed on the items and subscales making up the CEQ (Aleamoni, Note 2), for example, have yielded item reliabilities ranging from .73 to .94 and subscale reliabilities ranging from .80 to .98. It should be noted, however, that wherever student rating forms are not carefully constructed with the aid of professionals, as in the

case of most student-generated forms (Everly & Aleamoni, 1972), the reliabilities may be so low as completely to negate the evaluation effect and its results.

To answer the problem of the validity of student rating forms is much more difficult than addressing the reliability problem. To validate something one must have a criterion measure for comparison. One criterion measure that can be used to validate student rating forms is to determine how well the items and subscales measure what is intended (called content validity). This is usually accomplished by constructing the instrument so that it contains items and subscales that will yield measures in the areas that are considered necessary by an individual or group of experts in the field under consideration. Most of the student rating forms generated were validated by using this approach (Costin et al., 1971). Statistical tools like factor analysis have also been used to verify subjectively determined dimensions of the instructional setting and process. Both statistical (factor analysis) and subjective expert judgments were used in generating the items and subscales that make up the form (Aleamoni & Spencer, 1973).

Additional evidence that students can accurately rate instructors was presented in a recent study by Subkoviak and Levin (1974) showing a high degree of similarity between students' perceptions of the "ideal professor" at two separate universities.

Many other criterion measures have been suggested to validate student ratings. Some of those are peer (or colleague ratings), expert judges' ratings, and student learning. Many of the problems and questions that faculty pose, such as the eight stated above, can also be interpreted as validity concerns. To avoid any redundancy of answers to the faculty concerns, let me indicate that studies in which student ratings were compared to (a) colleague rating (Aleamoni & Yimer, 1973; Guthrie, 1954; Swanson & Sisson, 1971), (b) expert judges' ratings (Stallings & Spencer, Note 3), and (c) student learning measures (Cohen & Berger, 1970) all indicated the existence of high positive correlations which can be considered as providing additional validity evidence.

6. What extraneous variables or conditions affect student ratings? Studies conducted on the (a) size of the class (Aleamoni & Graham, 1974; Costin et al., 1971; Guthrie, 1954), (b) sex of the student (Costin et al., 1971; Aleamoni, Note 2), (c) time of day the course was offered (Aleamoni, Note 2), and (d) term (or semester) the course was offered (Costin et al., 1971; Aleamoni, Note 2), indicate that these variables had little or no relationship to the student ratings. The rank of the instructor (Aleamoni & Graham, 1974; Costin et al., 1971; Guthrie, 1954) seems to have some effect, but it is usually not statistically significant.

One the other hand, (a) whether the student was taking the course as a requirement or on an elective basis (Costin et al., 1971; Gillmore & Brandenburg, Note 6), and (b) the level of the course (Aleamoni & Graham,

1974; Costin et al., 1971) yielded significant effects on student ratings. Such effects, however, can be controlled through the use of appropriate normative data; this is an important feature of the results reported for the CEQ.

The one study cited by Costin et al. (1971) indicated that whether the students were majors or nonmajors did not affect their faculty ratings. This concern is presently being investigated using CEQ data.

7. The grades or marks students receive in the course are highly correlated to their ratings of the course and the instructor. There is ample evidence in the research literature to provide a definite answer to this concern. In almost all of the studies cited in Costin et al. (1971) and by investigators such as Guthrie (1954), Remmers (1960), and Weaver (1960), little or no relationship has been found between a student's grade and faculty rating. In fact, the positive correlations seldom exceed .30. The evidence, therefore, indicates that students do not necessarily rate an instructor or course based upon the grade they have or are about to receive.

8. How can student evaluations possibly be used to improve instruction? This could well be the most important question asked concerning student evaluation of instruction. There has been a great deal of discussion in the research literature about how, when, and where such evaluations should be used, but no clear-cut evidence has been offered to show that it does have an effect on instruction. The studies by Braunstein, Klein, and Pachla (1973); Centra (1973); and Miller (1971) were inconclusive with respect to the effect of feedback at mid-term to instructors whose instruction was again evaluated at the end of the term. However, such evidence is found in a recently completed study by Aleamoni (Note 7) where student ratings gathered near the end of the term using the CEQ were presented to university instructors along with the opportunity to discuss their results with a measurement and evaluation expert. The CEQ ratings were again gathered at the end of the same courses (one semester to a year later) taught by the same instructors. The results revealed that there was a significant increase in the student ratings of these faculty on the two lowest-rated CEQ subscales that were discussed in the meetings with the measurement and evaluation expert. On the other hand, faculty who were not able to avail themselves of expert consultation but did receive the CEQ results remained unchanged in their subscale ratings. Even though this study should be replicated, it represents the first tangible evidence that student ratings can be used to improve instruction.

It should be obvious by now that the problems and questions faculty typically raise about the appropriateness of using student ratings of instructor and instruction have very definite answers highly supportive of using such ratings. Interestingly enough, many faculty will still disregard the evidence and maintain that there is not enough evidence or that there are still too many unanswered questions and problems to take student ratings

seriously. However, these same faculty would stoutly defend their own methods of evaluating students even though they might not be able to present any evidence to substantiate their claims.

Conclusion

In conclusion, therefore, it might be useful to reverse the situation and ask faculty members to consider the eight concerns posed above from the point of view of **faculty evaluating students!** How much evidence would or could be provided to convince the students that:

1. Faculty can consistently judge student learning?
2. The performance of students with excellent ability is not used to set the standard for the rest of the students in the course?
3. Most faculty grading schemes are not affected by the attentive, polite, conforming, and noncreative student?
4. What was taught in the course is useful in other courses or outside that university?
5. The course examinations are reliable and valid?
6. An instructor's marks or grades are not affected by (a) the size of the class? (b) the sex of the student? (c) the time of day the course was offered? (d) whether the student was taking the course as a requirement or an elective? (e) whether the student was a major or a non-major? (f) the term (or semester) the course was offered? (g) the level of the course? (h) his/her professional rank?
7. Instructors who had a particularly rough time when they were in college do not tend to be just as rough on their students?
8. The course examinations are useful in improving their learning?

Cited Notes

1. Aleamoni, L. M. *Illinois course evaluation questionnaire (CEQ) results interpretation manual form 66 and form 32* (RR 331). Urbana, Il.: University of Illinois at Urbana-Champaign, Measurement and Research Division of the Office of Instructional Resources, 1972.
2. Aleamoni, L. M. *A review of recent reliability and validity studies on the Illinois course evaluation questionnaire (CEQ)* (RM 127). Urbana, Il.: University of Illinois at Urbana-Champaign, Measurement and Research Division of the Office of Instructional Resources, 1972.
3. Stallings, W. M., & Spencer, R. E. *Ratings of instructors in Accountancy 101 from video-tape clips* (RR 265). Urbana, Il.: University of Illinois at Urbana-Champaign, Measurement and Research Division of the Office of Instructional Resources, 1967.
4. Feldman, S. *The advisor* (RR 325). Urbana, Il.: University of Illinois at Urbana-Champaign, Measurement and Research Division of the Office of Instructional Resources, 1970.
5. Aleamoni, L. M., & Yimer, M. *Graduating senior ratings relationship to colleague rating, student rating, research productivity and academic rank in rating instructional effectiveness* (RR 352). Urbana, Il.: University of Illinois at Urbana-Champaign, Measurement and Research Division of the Office of Instructional Resources, 1974.
6. Gillmore, G. M., & Brandenburg, D. C. *Would the proportion of students taking a class as a requirement affect the student rating of the course?* (RR 347). Urbana, Il.: University of Illinois at Urbana-Champaign, Measurement and Research Division of the Office of Instructional Resources, 1974.

7. Aleamoni, L. M. *The usefulness of student evaluations in improving college teaching* (RR 356). Urbana, Il.: University of Illinois at Urbana-Champaign, Measurement and Research Division of the Office of Instructional Resources, 1974.



BOOK REVIEWS

References

- Aleamoni, L. M. Response to Professor W. Edward Deming's "memorandum on teaching." *The American Statistician*, 1972, 26 (4), 54.
- Aleamoni, L. M., and Graham, M. H. The relationship between CEQ ratings and instructor's rank, class size and course level. *Journal of Educational Measurement*, 1974, 11(3), 189-202.
- Aleamoni, L. M., and Spencer, R. E. The Illinois course evaluation questionnaire: A description of its development and a report of some of its results. *Educational and Psychological Measurement*, 1973, 33, 669-684.
- Aleamoni, L. M., and Yimer, M. An investigation of the relationship between colleague rating, student rating, research productivity, and academic rank in rating instructional effectiveness. *Journal of Educational Psychology*, 1973, 64, 274-277.
- Braunstein, D. N., Klein, G. A., and Pachla, M. Feedback expectancy and shifts in student ratings of college faculty. *Journal of Applied Psychology*, 1973, 58, 254-258.
- Centra, J. A. Effectiveness of student feedback in modifying college instruction. *Journal of Educational Psychology*, 1973, 65, 395-401.
- Cohen, S. H., and Berger, W. G. Dimensions of student's ratings of college instructors underlying subsequent achievement on course examinations. *Proceedings of the 78th Annual Convention of the American Psychological Association*, 1970, 5, 605-606.
- Costin, F., Greenough, W. T., and Menges, R. J. Student ratings of college teaching: Reliability, validity, and usefulness. *Review of Educational Research*, 1971, 41, 511-535.
- Deming, W. E. Memorandum on teaching. *The American Statistician*, 1972, 26(1), 47.
- Drucker, A. J., and Remmers, H. H. Do alumni and students differ in their attitudes toward instructors? *Purdue University Studies in Higher Education*, 1950, 70, 62-64.
- Drucker, A. J., and Remmers, H. H. Do alumni and students differ in their attitudes toward instructors? *Journal of Educational Psychology*, 1951, 42, 129-143.
- Everly, J. C., and Aleamoni, L. M. The rise and fall of the advisor. . . students attempt to evaluate their instructors. *Journal of the National Association of Colleges and Teachers of Agriculture*, 1972, 16, 43-45.
- Grush, J. E., and Costin, F. The student as consumer of the teaching process. *American Educational Research Journal*, 1975, 12, 55-66.
- Guthrie, E. R. *The evaluation of teaching: A progress report*. Seattle, Washington: University of Washington, 1954.
- McKeachie, W. J. Student ratings of faculty. *American Association of University Professors Bulletin*, 1969, 55, 439-444.
- Miller, M. T. Instructor attitudes toward, and their use of, student ratings of teachers. *Journal of Educational Psychology*, 1971, 62, 235-239.
- Remmers, H. H. *Manual of instruction for the Purdue rating scale for instruction*. West Lafayette, Indiana: University Book Store, 1960.
- Subkoviak, M. J., and Levin, J. R. Determining the characteristics of the ideal professor: An alternative approach. *Journal of Educational Measurement*, 1974, 11, 269-276.
- Swanson, R. A., and Sisson, D. J. The development, evaluation, and utilization of a departmental faculty appraisal system. *Journal of Industrial Teacher Education*, 1971, 9, 64-79.
- Weaver, C. H. Instructor rating by college students. *Journal of Educational Psychology*, 1960, 51, 21-25.

A. W. Burger
Book Review Editor
Dept. of Agronomy
Univ. of Illinois

John R. Campbell and Robert T. Marshall. **The Science of Providing Milk for Man**. McGraw-Hill Book Company, 1975, 801 pp. Hardbound \$16.95.

The title of this book indicates it will differ from the routine. Reading the book closely, one finds that it differs from many previous textbooks on this subject because it excellently covers the subject of milk from the cow completely through the finished product. A real attempt was made to relate milk to man.

The content is current. Sire summaries, production, testing programs, world and U.S. production and consumption, milking facilities and equipment, milk processing and milk products are thoroughly discussed and well illustrated. Some improvement could have been made in selecting pictures relating to manufacturing plants.

The strong points for this book include its depth of coverage, currency of information, quotations of famous people and historical times throughout the book, its sound scientific basis for all subjects and the easiness of reading. Possibly more attention was needed for the influence of environmental regulatory agencies on the dairy industry and some photographs should have been more current.

The book can be used very successfully in an introductory course covering both production and manufacturing, or it could be used in more advanced courses in either subject matter area. In addition, it certainly is an excellent reference source for anyone interested in knowing more about the dairy industry. I believe this book will receive widespread usage among our colleges and universities as a textbook.

Gary T. Lane
Dairy Production
Texas A & M University