DOCUMENT RESUME

ED 088 942 TM 003 503

AUTHOR Villano, Maurice W.; Rosenstock, Edward H.
TITLE A Decade with the Course Attitude Questionnaire: A Factorial Study.
PUB DATE Nov 73
NOTE 18p.; Paper presented at the Annual Convocation, Northeastern Educational Research Association (4th, Ellenville, NY, November 2, 1973)

EDRS PRICE MF-$0.75 HC-$1.50
DESCRIPTORS *Course Evaluation; Educational Diagnosis; *Evaluation; Factor Analysis; Item Analysis; *Longitudinal Studies; Norms; Participant Satisfaction; *Questionnaires; Reliability; *Student Attitudes; Validity
IDENTIFIERS *Course Attitude Questionnaire (CAQ)

ABSTRACT

The Course Attitude Questionnaire (CAQ) has been administered 100,000 times in 3,500 classes at the Pennsylvania State University since 1964 with extensive use in modified form at other institutions. The purpose of this research was to investigate the reliability and internal validity of the CAQ and the temporal stability of its original factor structure. CAO responses were obtained from 5,297 Penn State students in 272 classes during a 1973 term. These data were item-analyzed and submitted to principal-components and common-factor analyses with orthogonal and oblique transformations. Coefficient alpha reliabilities were computed for the factor-derived scales. The large normative data base was also inspected. The factor structure was not replicated by a "fit' of the derived factors to the CAQ-keyed subscales. Most factors were highly correlated and their reliabilities appeared to be a function of redundant item content. A "general" dimension pervaded the instrument, accounting for much of the total and common-factor variance. A trend toward inflation and skewness in scores was noted for the 10-year period. The CAQ) subscales should be interpreted with caution and systematically reviewed for currant validity. The CAQ, originally included in the Appendix of this document, has been removed because it does not reproduce. For related document see ED 029508. (Author/RC)

A DECADE WITH THE COURSE ATTITUDE QUESTIONNAIRE:

A FACTORIAL STUDY

Maurice W. Villano and Edward H. Rosenstock

The Pennsylvania State University

(The Course Attitude Questionnaire, originally included in the Appendix of this document, has been removed because it does not reproduce.)

Paper presented at the Fourth Annual Convocation
Northeastern Educational Research Association
Ellenville, New York November 2, 1973

A DECADE WITH THE COURSE ATTITUDE QUESTIONNAIRE: A FACTORIAL STUDY

Maurice W. Villano and Edward H. Rosenstock

The Pennsylvania State University

Background

The Course Attitude Questionnaire (CAQ) of R. E. Spencer has been administered over 100,000 times in almost 3,500 classes at the Pennsylvania State University since 1964. The instrument was originally developed for instructional television but was later generalized to include all aspects of instruction as a basis for course improvement.(Dick, 1965). The CAQ has also been used extensively in its present or modified form at other institutions and a large normative data base has been established.

There has been a decline, however, in the use of the CAQ at the Pennsylvania State University. Over half of the major colleges or departments and several of the 18 Commonwealth branch campuses of the University now favor other rating forms adopted from nationally-known instruments or constructed locally. Criticisms by students and faculty have alleged that the Course Attitude Questionnaire is "too broad" for evaluative or diagnostic purposes, the item statements are repetitive or overlapping, and some areas are not represented.

The feasibility of continued use of the CAQ on a University-wide basis has been under consideration. The purpose of this paper was to report an investigation of the current reliability and validity of the Course Attitude Questionnaire, and the temporal stability of its originally-claimed factor structure. A large sample of contemporary data was used in an attempt to replicate the earlier research of Dick (1965) and Spencer and Aleamoni (1969).

Research has not kept pace with the development and use of student evaluation instruments (Bejar, 1973) and findings reported in the literature have been characterized as inconsistent and often contradictory in their nature. The decrease in the use of rating forms, such as the CAQ, may be due to a lack of convincing validation data (Costin, Greenough & Menges, 1971). One of the most recent and extensive summaries in this area of concern was the First Invitational Conference on Faculty Effectiveness as Evaluated by Students held at Temple University earlier this year (Sockloff, 1973). Among some of the prominent discussants were J. Centra of ETS, P. Dressel and W. Warrington of Michigan State University, W. McKeachie of the University of Michigan, and R. Perry and R. Baumann of the University of Toledo. A general consensus that emerged from this meeting reaffirmed the need for: (1) longitudinal studies; (2) instruments that were observable, ratable, and acceptable; and (3) scoring and scaling systems that adequately discriminated among instructors or courses.

There has been a dearth of long-range evaluations of some of the more-extensively used student rating forms. The findings of this investigation may therefore have some important educational implications. Despite the seeming limitations of instruments like the CAQ, there is sufficient convergent evidence to suggest possible directions for their improvement. Minimally, there is a requirement for the periodic and systematic review of student evaluation devices to assure their current validity.

Factorial studies, as in this investigation, have examined the reduced sets of dimensions that described and explained the variance of factors underlying student evaluations. Reliable factor-derived dimensions or clusters of items are sought in the analysis of the instrument and its sub-scales (Guertin & Bailey, 1970). Some approximate methods for comparing the temporal stability of factor structures are discussed by Nunnally (1967) and others.

Procedures

The instrument evaluated in this replication study was the Course Attitude Questionnaire. The CAQ can be described as having 50 forced-choice Likert-type items, including 22 matched positive-negative item-pairs, on a four-step Agree-Disagree scale (see Appendix). Negative items are reversed in scoring, and the cumulative, unweighted item responses are then summed to yield scores on each of six keyed sub-scales of General Course Attitude, Method, Content, Interest-Attention, Instructor, and "Other," as well as a Total CAQ score.

The response data source was provided by 7,031 Pennsylvania State University students administered the CAQ in 272 graduate and undergraduate classes during the Winter 1973 term. The sample was screened for missing item responses and reduced to 5,297 Ss.

The cumulative normative data compiled on the CAQ were reviewed and comparisons of descriptive statistics were made at selected points during the past 10 years. The reliability coefficients for single-form internal consistency (split-half method) were calculated for groups of positive and negative items and between groups of half-positive and half-negative items (Nunnally, 1967). The estimated Coefficient Alpha reliabilities were computed for both the CAQ-keyed and factor-derived sub-scales or item clusters (Cronbach, 1951).

At the item level, the response data were evaluated by multiple modes of factor analysis as suggested by Glass and Taylor (1966). The generated 50 by 50 intercorrelation matrix was submitted to a principal-components analysis, with unities in the diagonal, to replicate the derived dimensions and total test variance explained by the original CAQ factor structure (Hotelling, 1933). Some limitations of this approach are explained at the end of the section. The "R" matrix was also subjected to an Alpha factor analysis, with estimates of

communalities inserted in the diagonal, to make inferences regarding the generalizability of the derived factors to the CAQ item content domain and to account for the common-factor variance (Kaiser & Caffrey, 1965). A root criterion of unity was used for the extraction of factors in both analyses which were then submitted to orthogonal (Varimax) rotations (Kaiser, 1958) and oblique (Oblimin) transformations (Harman, 1967). The most psychologically-meaningful and psychometrically-defensible solutions were chosen for interpretation.

This study was somewhat limited by the nature of the information that was available from the earlier investigations (Dick, 1965; Spencer & Aleamoni, 1969). For example, the factor loading matrix was not reported thus precluding direct comparison as in the computation of coefficients of congruence for pairs of factor vectors across the CAQ data samples (Rummel, 1970). The basis for the assignment of the CAQ items to the sub-scales was brought into question (Dick, 1965):

> Listed below are the factors and the items which were assigned to each. The per cent of variance in the students' scores which is generally accounted for by each of the factors is also shown. These percentages are an approximate average which has been obtained from subsequent factor analyses of the questionnaire (p. 4). /underlining supplied/

Factorial invariance was apparently hinted at as well as the "contrived" structure of the "Other" sub-scale (Ibid.):

> The factor analyses of the responses from various courses indicate that the factors identified above are fairly stable; with the exception of some of the items which are related to course content, the items generally cluster together as indicated above. Some of the course content items occassionally (sic) intercorrelate quite highly with the more general course attitude items (Factor 1).
>
> . . . . . .
>
> A final note should be made of the items in Factor 6 - specific or other items. Although these items did not inter-correlate with the other factors, they were retained in the present form of the questionnaire because of the specific information which is available in the responses to these items. Although the responses to these items are grouped together to form one factor score, it is possible to get the individual distributions of responses to these and other items on the questionnaire (pp. 7-8). /underlining supplied/

Findings

There was a discernable inflationary trend in Course Attitude Questionnaire ratings over the past decade. General increases in the CAQ sub-scale and total scores were noted for the time intervals indicated (see Table 1). An inspection of the cumulative norms through 1973 revealed there was not a sufficient spread in the class means for the courses to discriminate adequately except at the extreme points (10th- and 90th-percentiles) of the distribution (see Table 2).

The substantial reliabilities (.80's and .90's) for the CAQ-defined or keyed sub-scales and the overall instrument that had been reported in the earlier studies were confirmed (see Table 3). The magnitude of these reliability coefficients was due in part to the possibly spurious effect of the redundant item-pairs that contributed to the high intercorrelations of items within scales.

The obliquely-rotated seven-factor soulutions for both the principal-components and Alpha factor analyses were determined to most closely approximate the purported CAQ factor structure. The results of the two modes factor analysis were similar except for minor differences in the factor assignments of a few items having "weak" loadings (4, 7, 32, and 42); also, the loading coefficients of the Alpha factors were generally lower (see Tables 4 and 5). The coefficients of congruence accross both factoring methods approached unity signifying almost perfect agreement.

The empirically-derived factor scales did not provide a "close fit" to the a priori or CAQ-defined or keyed sub-scales. The artificiality" of the "Other" sub-scale was verified by its fission into two factors (V and VII). One-fourth of the items loaded highest on "non-relevant" factors. Several items exhibited a factorial complexity of two or more. Some items failed to achieve loading coefficients of .40 or greater on any factor.

Some weak Alpha factors suggested that the sampling from the content universe of CAQ items may not have been complete. Fewer items clustered on the derived factors that resembled the Content, Interest-Attention, and "Other" sub-scales as compared to those for General Course Attitude, Method, and Instructor (see Table 5). The reliabilities for the former group were also somewhat lower (see Table 6). For these factors, at least, a reliable inference could not be made with respect to their "positive generalizability" to the content domain from which the items had been drawn. The shift of items from their original CAQ-keyed assignments had been observed (see Table 7).

A "general" dimension appeared to pervade the overall CAQ instrument. The first extracted factor in both analyses accounted alone for 72-percent of the common variance and 43-percent of the total variance. Furthermore, the primary axes of the factors corresponding to the General Course Attitude, Method, and Interest-Attention sub-scales were highly intercorrelated (above .50) (see Table 8). It was concluded that the factorial invariance of the Course Attitude Questionnaire had not been demonstrated. It was recommended that scores derived from the CAQ sub-scales be interpreted with caution and that research on the relationship of the CAQ with important educational variables proceed with care.

It was further recommended that any revised version of the CAQ be patterned after the model of the current Illinois Course Evaluation Questionnaire (CEQ Form 72) where duplicative items have been eliminated or rewritten, optional items incorporated for specific diagnostics, and provision made for open-ended or free-response comments by students (Aleamoni, 1973).

References

Aleamoni, L. The usefulness of student evaluations in improving college teaching. In A. Sockloff (Ed.) Proceedings: First Invitational Conference on Faculty Effectiveness as Evaluated by Students. Philadelphia: Temple University Measurement and Research Center, 1973. (In press)

Bejar, I. A survey of selected administrative practices supporting student evaluation of instruction programs. Minneapolis: University of Minnesota Measurement Services Center, 1973. 23 pp.

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

Cronbach, L. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16, 297-334.

Dick, W. Course Attitude Questionnaire: its development, uses, and research results. Report No. 106, Office of Examination Services, University Division of Instructional Services. University Park, Pa.: The Pennsylvania State University, 1965. 21 pp.

Glass, G., & Taylor, P. Factor analytic methodology. Review of Educational Research, 1966, 36. 566-587.

Guertin, W., & Bailey, J. Introduction to Modern Factor Analysis. Ann Arbor, Mich.: Edward Bros., Inc., 1970.

Harman, H. Modern Factor Analysis. Chicago: University of Chicago Press, 1967.

Hotelling, H. Analysis of a complex set of statistical variables into principal components. Journal of Educational Psychology, 1933, 24, 417-441, 498-520.

Kaiser, H. The varimax criterion for analytic rotation in factor analysis. Psychometrika, 1958, 23(3), 187-200.

Kaiser, H., & Caffrey, J. Alpha factor analysis. Psychometrika, 1965, 30(1), 1-14.

Nunnally, J. Psychometric Theory. New York: McGraw-Hill, 1967.

Rummel, R. Applied Factor Analysis. Evanston, Ill.: Northwestern University Press, 1970.

Sockloff, A. (Ed.) Proceedings: First Invitational Conference on Faculty Effectiveness as Evaluated by Students. Philadelphia: Temple University Measurement and Research Center, 1973. 160 pp. (In press)

Spencer, R., & Aleamoni, L. 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. 17 pp.

TABLE 1

SELECTED CUMULATIVE NORMATIVE DATA FOR THE CAQ AT THE PENNSYLVANIA STATE UNIVERSITY: 1964 - 1973[a]

Year[b]	No. of Classes	Gen Crse Attitude	Method	Content	Interest-Attention	Instructor	Other	Total
1964	115	24.5	20.0	22.5	20.7	24.4	28.3	140.3
1966	235	24.5	21.3	22.5	21.2	24.7	28.6	142.8
1969	723	24.0	22.0	22.7	22.8	25.7	28.1	146.6
1970	1096	24.8	22.9	23.0	22.7	25.3	28.8	147.6
1971	1371	24.6	22.8	22.7	22.7	25.3	29.0	147.5
1972	2896	24.6	23.0	23.0	22.6	25.6	28.7	148.3
1973	3455	25.2	23.0	23.3	23.0	25.2	29.4	148.6
1973[c]	3455	(24.7)	(22.9)	(23.1)	(22.7)	(25.4)	(29.0)	(147.8)

[a]Norms are for scores at the 50th-percentile (median).

[b]Last academic term of calendar year for which data are available.

[c]Actual means indicated in parentheses.

TABLE 2

CUMULATIVE NORMATIVE DATA FOR THE COURSE ATTITUDE QUESTIONNAIRE

AT THE PENNSYLVANIA STATE UNIVERSITY THROUGH SUMMER TERM 1973

	Percentile						
Sub-scale	10	25	50	75	90	$\bar{X}$	SD
Gen Crse Attitude	21.7	22.9	25.2	26.8	28.0	24.7	2.5
Method	19.2	20.8	23.0	25.0	27.0	22.9	3.0
Content	20.6	22.2	23.3	24.7	25.8	23.1	2.0
Interest-Attention	18.6	20.9	23.0	24.6	26.1	22.7	3.0
Instructor	22.7	24.1	25.2	27.0	28.1	25.4	2.1
Other	26.6	27.6	29.4	30.0	31.8	29.0	2.1
Total	129.8	139.5	148.6	156.9	165.3	147.8	13.8

TABLE 3A

SPLIT-HALF RELIABILITIES FOR THE CAQ

Groups of Items	1969 Study[a] (N = 297)	1973 Study (N = 5,297)
22 positive versus 22 negative matched-pair items	.92	.92
50 items in two groups of half-positive and half-negative items	.93	.91

[a]Spencer & Aleamoni (1969, pp. 7-8)

TABLE 3B

COEFFICIENT ALPHA RELIABILITIES FOR THE CAQ-KEYED SUB-SCALES

Sub-Scale	No. of Items	1965 Study[a] (N = 460)	1973 Study (N = 5,297)
Gen Crse Attitude	8	.85	.93
Method	8	.92	.94
Content	8	.56	.81
Interest-Attention	8	.89	.94
Instructor	8	.72	.82
Other	10	.57	.75

[a]Dick (1965, p. 14)

TABLE 4

PRIMARY PATTERN MATRIX FOR DERIVED CAQ COMPONENTS SOLUTION USING OBLIMIN TRANSFORMATION[a]

CAQ-Key	Item No.	Gen Crse Attitude I	Method IV	Content II	Interest-Attention VI	Instructor III	Other-1 V	Other-2 VII
C	13	75						
G	49	73						
G	20	73						
C	19	71						
G	25	70						
G	34	69						
G	11	65						
IA	24	59						
C	40	57						
IA	35	56			32			
G	2	54						
G	3	49	35					
IA	7	46			40			
G	29	41		35				
M	37		87					
M	8		86					
M	48		84					
M	1		82					
M	27		73					
M	6		71					
M	50		63					
M	36		53			35		
O	42		32					
C	39			70				
C	28			69				
C	30			66				
O	43			63				
O	41			52				39
C	26			35				

TABLE 4 (Continued)

CAQ-Key	Item No.	Gen Crse Attitude I	Method IV	Content II	Interest-Attention VI	Instructor III	Other-1 V	Other-2 VII
IA	9				62			
IA	14				59			
IA	45				58			
IA	22	34			51			
IA	46	41			42			
O	17				35			
IN	47					70		
IN	5					69		
IN	31					69		
IN	18					60		
IN	12					53		
IN	23					49		
IN	10		30			44		
O	33					32		
O	38						72	
C	44			31			67	
O	32			45			56	
O	16							62
O	21							50
IN	15							44
O	4	37						40
Percentage of total variance accounted for:		42.5%	3.3%	5.1%	2.4%	4.5%	2.5%	2.2%

[a]Loadings have been multiplied by 100 and rescaled to eliminate negative signs. Loadings less than .30 deleted.

TABLE 5

PRIMARY PATTERN MATRIX FOR DERIVED CAQ ALPHA FACTOR SOLUTION USING OBLIMIN TRANSFORMATION[a]

CAQ-Key	Item No.	Gen Crse Attitude I	Method IV	Content II	Interest-Attention VI	Instructor III	Other-1 V	Other-2 VII
C	13	70						
G	49	70						
C	19	69						
G	25	66						
G	20	66						
G	34	65						
G	11	59						
C	40	55						
IA	24	53						
IA	35	52			33			
G	2	48						
G	3	46	33					
G	29	36		35				
O	4	32						
O	42	31						
M	37		85					
M	8		83					
M	48		81					
M	27		70					
M	1		69					
M	6		68					
M	50		60					
M	36		50			38		
C	39			62				
C	28			60				
C	30			59				
O	43			55				
O	41			44				36
O	32			33				
C	26			31				

TABLE 5 (Continued)

CAQ-Key	Item No.	Gen Crse Attitude I	Method IV	Content II	Interest-Attention VI	Instructor III	Other-1 V	Other-2 VII
IA	9				61			
IA	14				59			
IA	22				50			
IA	45				47			
IA	7	40			41			
IA	46	34			41			
O	17				19			
IN	5					65		
IN	31					63		
IN	47					56		
IN	18					52		
IN	12					46		
IN	23					45		
IN	10					42		
O	33					31		
O	38						69	
C	44						62	
O	21							43
O	16							42
IN	15							27
Percentage of common variance accounted for:		72.0%	4.5%	8.0%	2.7%	7.1%	3.3%	2.6%

[a]Loadings multiplied by 100 and rescaled to eliminate negative signs. Loadings less than .30 deleted (except primary loadings on Items 15 and 17).

TABLE 6

COEFFICIENT ALPHA RELIABILITIES

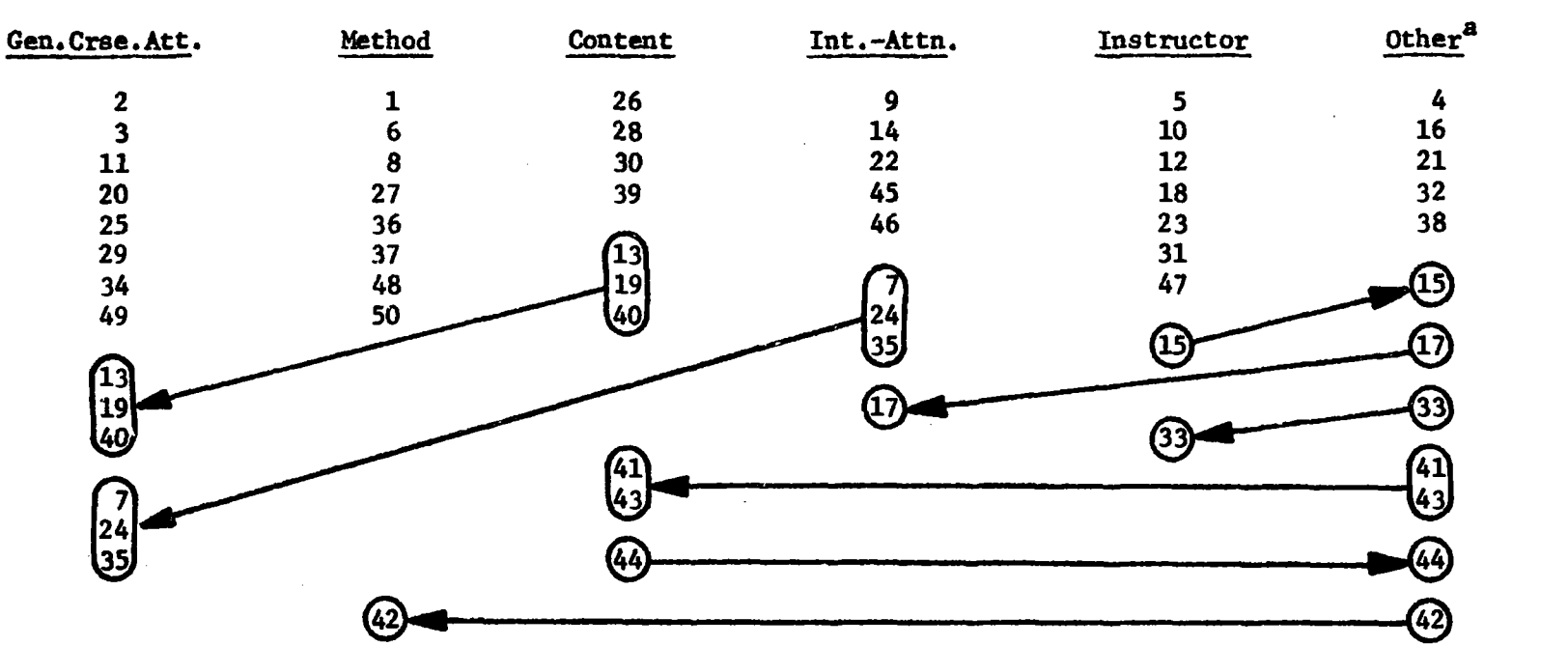
FOR THE CAQ FACTOR-DERIVED SUB-SCALES

Factor[a] Scale	Corresponding CAQ Sub-scale	Components Reliability	No. of Items	Alpha Factor Reliability	No. of Items
I	Gen.Crse.Att.	.96	14	.95	15
IV	Method	.94	9	.94	8
II	Content	.83	6	.81	7
VI	Int.-Attn.	.89	6	.91	7
III	Instructor	.83	8	.83	8
V & VII	Other	.59	7	.62	5

[a]Two factors resembling the "Other" sub-scale were coalesced for ease of interpretation.

TABLE 7

SHIFT OF ITEMS IN CHANGE OF CAQ COMPONENT STRUCTURE



[a]Factors V (Other-1) and VII (Other-2) of the derived solution were combined for ease of interpretation.

TABLE 8

INTERCORRELATION OF THE PRIMARY AXES FOR THE CAQ COMPONENTS AND ALPHA FACTOR SOLUTIONS[a]

	Component or Factor[b]					
Component or Factor	I Gen.Crse.Att.	IV Method	II Content	VI Int.-Attn.	III Instructor	V Other-1
IV	55(57)					
II	-27(-32)	-41(-49)				
VI	36(47)	44(51)	-22(-29)			
III	-38(-45)	-49(-57)	22(30)	-26(-32)		
V	-18(26)	-24(30)	03(-02)	-09(12)	23(-33)	
VII Other-2	-26(31)	-27(28)	16(-22)	-20(24)	26(-33)	04(00)

[a]Alpha factor intercorrelations are in parentheses.

[b]All entries have been multiplied by 100.