DOCUMENT RESUME

ED 055 154 VT 010 947

AUTHOR Taylor, C. W.; And Others
TITLE Synthesis of Multiple Criteria of Physician Performance.
INSTITUTION Utah Univ., Salt Lake City. Dept. of Psychology.
SPONS AGENCY Public Health Service (DHEW), Washington, D.C. Bureau of Health Manpower.
PUB DATE Nov 68
NOTE 15p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Evaluation Criteria; *Evaluation Methods; Evaluation Techniques; Manpower Development; *Performance Criteria; *Physicians; *Statistical Analysis; Synthesis

ABSTRACT

In order to evaluate on-the-job physician performance, four different composite scores were derived from different weightings of 80 criteria. Weights were assigned both objectively, through statistical methods, and subjectively, based only upon personal judgment. Comparing these results with scores based upon appraisal of the individual by other physicians produced remarkably similar results, especially considering the instability of many of the criteria. (BH)

ED055154

128 PM00017

PM 00017

Abstract No. 42: Program: Seventh Annual Conference on Research in Medical Education, 1968, pp. 30-31.
The Association of American Medical Colleges

Syntheses of Multiple Criteria of Physician Performance. C. W. Taylor, Ph.D., E. G. Lewis, Ph.D., D. E. Nelson, B.S., G. C. Loughmiller, M.A., and P. B. Price, M.D., University of Utah.

In 1964, Price, et al. (J. Med. Educ., 39: 203-211, 1964) reported a medical criterion study in which 80 criteria of physician performance were collected and analyzed for a sample of several hundred Utah physicians. The purpose of this paper is to determine how well these same criteria can be combined or synthesized into alternate versions of a single summary score of over-all physician performance. These summary scores provide new criteria for validating predictors. They will also serve as "landmarks" against which new criterion scores, soon to be collected, will be compared to see to what degree they extend the range of physician performances already measured.

Four different composite or summary scores for each of 205 specialists studies by Price, et al., were derived by weighting each of the 80 criterion measures in four different ways. (a) Composite scores were obtained by simply summing the 80 unweighted criterion scores for each physician. (b) The 80 criteria were summed after each criterion had been assigned a weight based on its judged importance to medical practice. These weights were assigned to the 80 criteria by a panel of five physician judges. (c) Medical experts gave an over-all rating of 1 to 10 to each specialist on the basis of his 80 performance criteria contained in a coded, anonymous folder. (d) The weight for each criterion was derived by reversing the typical application of the statistical technique of multiple correlation, using the 80 performance variables as "predictors" and the over-all rating of performance (described in c above) as the criterion.

After derivation of the composite criterion scores, half of the intercorrelations among the four synthesized criterion scores were found to be over .85.

We plan to obtain a fifth summary score by having a number of knowledgeable physicians make an over-all appraisal of each specialist by name. These ratings, obtained independently of the 80 criteria, will provide another indication of how well our syntheses of the 80 performance measures reflect actual physician practice. Early indications are that our composite scores hold much promise as reasonably accurate criteria of on-the-job physician performance for use until they can be further extended and refined.

AT010947

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

PM00017

Syntheses of Multiple Criteria of Physician Performance[1]

Taylor, C. W., Lewis, E. G., Nelson, D. E.
Loughmiller, G. C., and Price, P. B.

College of Medicine and Department of Psychology
University of Utah

In 1964, Price, et. al. (J. Med. Educ., 39:203-211, 1964) reported a medical criterion study in which 80 criteria of physician performance were collected and analyzed for a sample of several hundred Utah Physicians (see Table 1). In the first phase of the present study those 80 criteria were combined in various ways to produce four summary scores of overall physician performance. This procedure was deemed desirable for two reasons: (1) Such summary scores would provide additional criteria for validating predictors; (2) the summary scores could serve as "landmarks" against which new criterion scores, soon to be collected, might be compared. The degree to which these future criteria would extend the range of the summary performance measures could then be evaluated. The present study was accomplished on only our specialist sample: general practitioners and medical faculty were not included.

Method and Results

Before combining the 80 criterion scores for each physician, an inspection of the criteria revealed that a high score on some of the criteria did not necessarily indicate a high quality of performance. For example, an excessive number of consultations requested by a physician would increase a total summed score but might reflect an undesirable method of practice.

---

[1]This project was part of a larger project "Measurement and Prediction of Physician Performance" supported by the Department of Health, Education and Welfare, contract No. PM-00017-02

Table 1

Title and Source for Each Variable for Urban Specialist Sample

Title	Source
1. Number of Times Nominated as Outstanding Contributor by Urban Specialist Colleagues	Colleagues
2. Number of Times Nominated as Preferred Consultant by Urban Specialist Colleagues	Colleagues
3. Number of Nominations as Outstanding Contributor or Preferred Consultant by General Practitioners	Interview
4. Number of Times Nominated as Outstanding Contributor by College of Medicine Faculty	Interview
5. Rating of "Clinical Excellence" by Medical College Department Head	Department Head
6. Number of Different Residency Hospitals	Compendiums
7. Number of Years Spent in Residency	Compendiums
8. Judged Quality of National Board Certification	Expert Judges
9. Present College of Medicine Clinical Faculty Rank	Official Records
10. Mobility Rate in Professional Positions Since Receiving M.D.	Interview and Official Records
11. Total Number of Listings in Honorary Compendiums	Compendiums
12. Gross Income from Medical Profession	Interview
13. Number of Current Memberships in Scientific and Professional Societies	Interview
14. Average Judged Quality of Societies in Which Membership is Current	Interview and Expert Judges
15. Overall Occupational Satisfaction	Questionnaire
16. Number of Times During Career Invited to Serve as Editor of Scientific or Professional Journal	Interview
17. Number of Times During Career Invited to Serve on Scientific and Professional Advisory Boards	Interview
18. Average Judged Quality of Scientific and Professional Awards Received During Career	Interview and Expert Judges
19. Self-Reported Number of Contributions Made to Medicine	Interview
20. Average Judged Quality of Self-Reported Contributions Made to Medicine	Interview and Expert Judges
21. Self-Reported Number of Non-Medical Contributions to Society	Interview
22. Total Number of Papers Presented at Scientific and Professional Meetings During Career	Interview
23. Average Number of Journal Publications Per Year Since Receiving M.D.	Interview and Compendiums
24. Average Level of Contribution to Publications as Indicated by Senior vs. Junior Authorship Status	Interview and Compendiums
25. Number of Research Projects with Which Involved During Career	Interview
26. Number of Scientific and Professional Journals Reviewed Regularly	Interview
27. Number of Subscriptions to Scientific and Professional Journals	Interview
28. Number of Articles in Scientific and Professional Journals Read in Detail Each Month	Interview
29. Average Number of Society Meetings Attended Annually	Interview
30. Number of Postgraduate Courses Taken During Career	Interview
31. Number of Refresher Courses Taken During Career	Interview
32. Physician's Evaluation of Usefullness of Drug Detail Men	Interview
33. Extent of Physician's Experimental Use of Drugs Provided by Drug Detail Men	Interview
34. Number of Techniques Other Than Journals, Meetings, and Drug Detail Men Used in Keeping Abreast	Interview
35. Average Number of Formal Medical Consultations Called into Monthly	Interview
36. Average Number of Informal Medical Consultations Called into Monthly	Interview
37. Percentage of Patients on Which Consultations are Requested	Interview
38. Number of Patients Seen Per Day	Interview
39. Average Amount of Time Spent with Patients on First Visit	Interview
40. Average Amount of Time Spent in Explaining Diagnoses to Patients	Interview
41. Proportion of Office Patients Treated Without Charge	Interview
42. Proportion of Office Patients That Fail to Pay Physician for Services Rendered	Interview
43. Self-Estimated Average Socioeconomic Level of Patients	Interview
44. Degree to Which Physician Adheres to Patient Appointment Schedule	Interview
45. Average Number of House Calls Made Per Week	Interview
46. Degree to Which Physician Considers Psychological Factors in Diagnoses	Interview
47. Average Number of Hours Per Week Devoted to Medical Practice	Interview
48. Average Number of Hospitalized Patients	Interview
49. Number of Hospitals in Which Physician Works	Interview
50. Average Judged Quality of Hospitals in Which Physician Works	Interview and Expert Judges
51. Number of Hospitals in Which Physician Maintains Courtesy Priviledges	Interview
52. Average Judged Quality of Hospitals in Which Physician Maintains Courtesy Priviledges	Interview
53. Number of Formal Responsibilities Physician has in Hospitals	Interview
54. Average Judged Quality of Hospital Responsibilities	[illegible] Judges
55. Self-Estimated Value of Office Equipment	Interview
56. Number of M.D. Assistants on Physician's Ancillary Staff	Interview
57. Number of Nurses on Physician's Ancillary Staff	Interview
58. Number of Technicians on Physician's Ancillary Staff	Interview
59. Number of Clerical, Administrative, and Janitorial Workers on Physician's Ancillary Staff	Interview
60. Average Number of Speeches on Medical Topics to Laymen Groups Per Year	Interview
61. Average Amount of Vacation Taken Annually	Interview
62. Extent to Which Physician Plans and Maintains Leisure Time Activities	Interview
63. Number of Current Memberships in Social and Avocational Organization	Interview
64. Number of Current Memberships in Civic and Political Organizations	Interview
65. Characteristics Physician Considers Important for Success in Medicine: Number of "Common" Responses	Interview
66. Characteristics Physician Considers Important for Success in Medicine: Number of "Uncommon" Responses	Interview
67. Characteristics Physician Considers Important for Success in Medicine: "Commonness" of Responses Elicited	Interview
68. Self-Rating of Success in Medicine	Expert Judges
69. Expert Panel Rating of Overall Performance Based on All Available Information	Project Interviewer
70. Interviewer Rating of Condition of Physician's Office	Project Interviewer
71. Interviewer Rating of Likeability	Project Interviewer
72. Interviewer Rating of Physician's Involvement in This Project	Control Variables
73. Age at Which Received M.D.	Control Variables
74. Number of Years Between Receiving M.D. and Receiving National Board Certification	Control Variables
75. Years of Experience Since Receiving M.D.	Control Variables
76. Individual Practice Rather Than Group, Clinic or Hospital Practice	Control Variables
77. Hospital Practice Rather Than Individual, Group, or Clinic Practice	Official Transcripts
78. Undergraduate Grade Point Average	Official Transcripts
79. Grade Point Average for First Two Years of Medical School	Official Transcripts
80. Grade Point Average for Last Two Years of Medical School	Official Transcripts

Therefore, items on which a high score was not necessarily indicative of superior performance were submitted to a panel of five physician judges who estimated the optimum level for each of these criteria. Scores above this estimated optimum were prorated downward, making them comparable to lower scores on that criterion. Ten criteria were re-scored in this manner.

Four different summary scores were derived for each of the 205 specialists studied by Price, et. al., (1964). For the first three summary scores this was acomplished by weighting each of the 80 criterion measures in three different ways: (1) Judged Weights--the 80 criteria were summed after each criterion had been assigned a weight based on its judged importance to medical practice. These weights were assigned to the 80 criteria by having a panel of five physician judges rate the essentiality of each criterion for superior performance on a scale that ranged from 5 "essential" to 1 "inconsequential"; (2) Equal Weights--composite scores were obtained by simply summing the 80 unweighted criterion scores for each physician; and (3) Statistical Weights--the weight for each criterion was derived by reversing the typical applicatio of the statistical technique of multiple correlation, using the 80 performance variables as "predictors" and the direct overall rating of performance (described in 4 below) as the criterion. The beta-weights assigned to each variable in the regression equation were then used to weight the criteria. Every summary score mentioned above was achieved by weighting scores which had been previously standardized so that each criterion had essentially equal variability before weighting. The fourth score, an overall rating from a criterion folder

was obtained by having medical experts give an over-all rating of each specialist on the basis of his 80 performance criteria contained in his coded, anonymous folder. A rating of 1-10 was assigned to each physician, with lower scores indicating superior performance. After the four summary scores were derived for each of the 205 physicians, these scores were converted to percentile ranks, as shown in Table 2. From that table it can be noted that physicians received very similar percentile ranks for all four methods of score synthesis. This was true particularly at the high and low ends of the continuum

The next step in the present study was the securing of independent performance ratings on the physician sample, obtained independently of the data in each criterion folder. These appraisals were made by five knowledgeable physicians who were selected on the basis of having had a wide exposure to a large number of Utah physicians. Names of the 205 specialists in the sample under consideration were typed on 3 x 5 cards, one name to a card. After considering the quality of each physician's total performance, the judges sorted each name on a nine-point scale which ranged from 1 "outstanding" to 9 "extremely poor." A tenth category, "I cannot rate this physician" was also included. The five physician judges working independently of any knowledge of the previously collected criteria and their ratings were averaged. Column Five in Table 2 gives the percentile ranks derived from these average ratings. (Scores based on previously collected criteria are in columns 1-4).

The final step in the analysis of the five over-all performance measures consisted of intercorrelating these five variables. Table 3 gives the results of this intercorrelation. It can be seen that the

Table 2

FIVE OVERALL MEASURES OF PHYSICIAN PERFORMANCE

The first 3 columns below rank 205 specialists on 80 criterion measures of physician performance gathered to date. Each of these columns simply represents a different way of combining or synthesizing the 80 measures into one overall criterion of performance. Columns 4 and 5 are ratings assigned by physician judges. Each row represents the relative position of one physician, as ranked by the methods described briefly above and in more detail below:

1. Judged Weight. Measures were weighted according to essentiality for superior practice by 5 physicians and summed.
2. Equal Weight. All 80 performance measures were given an equal weight before summation.
3. Statistical Weight. Before summation, each measure we weighted by a multiple correlation which determined the relationship of each criterion to a global estimate of performance (see 4 below).
4. Overall rating of anonymous criterion folder. A 1-10 rating of overall performance was made by a panel of physicians who considered all 80 measures of each physician.
5. Overall rating of physician by name. A measure of overall performance independent of our criteria was obtained by having five physician judges rate each specialist on a 1-9 scale. A mean rating was then calculated for each specialist.

The above ratings and rankings were subsequently converted into the percentile ranks shown below:

PERCENTILE RANKS					PERCENTILE RANKS				
1	2	3	4	5	1	2	3	4	5
99	99	99	95	99	76	80	69	78	48
99	98	99	95	48	76	81	80	20	59
98	99	96	95	97	75	76	56	20	72
98	97	98	99	99	75	82	69	69	62
98	98	97	95	93	75	78	45	20	72
97	97	97	95	95	75	86	80	78	91
96	95	98	95	98	75	68	83	48	25
95	96	95	95	98	73	74	75	08	67
95	95	94	91	67	73	70	85	08	82
94	94	96	95	97	72	74	73	20	86
94	94	93	20	88	72	80	74	08	59
94	93	93	91	82	71	79	90	03	72
93	94	95	20	93	71	61	61	48	70
93	[illegible]	88	78	89	71	73	91	20	48
92	90	84	48	39	70	68	64	48	62
92	90	87	78	23	69	62	62	08	08
91	93	90	91	92	69	84	67	08	67
90	92	91	95	78	68	74	71	48	82
90	86	84	95	98	67	71	71	78	10
89	89	79	95	78	67	74	76	48	45
89	90	94	20	82	66	66	61	78	76
88	85	68	48	59	65	61	83	08	62
88	87	63	08	38	65	51	16		62
88	88	53	48	86	65	59	77	48	82
87	88	89	48	70	65	67	78	20	76
86	83	86	78	62	64	59	38	48	23
86	81	78	78	48	64	71	45	20	78
86	83	77	95	84	64	61	86	78	72
85	91	89	78	89	64	56	24	20	16
85	88	82	20	01	64	62	85	03	86
85	82	70	20	25	63	71	57	20	83
84	78	66	08	95	63	63	56	08	78
84	85	87	48	16	62	69	60	08	93
83	77	64	48	23	62	71	67	48	82
82	84	65	20	16	61	55	46	20	89
81	73	92	01	78	61	39	15	20	62
80	85	68	20	67	61	68	52	20	02
80	78	79	03	76	60	69	50	48	28
79	86	92	20	76	60	60	51	20	32
79	83	62	03	82	60	51	51	48	82
78	75	74	48	48	59	55	57	78	48
77	71	86	48	67	58	57	34	20	65
76	82	81	78	48	57	48	55	48	95

PERCENTILE RANKS

1	2	3	4	5
57	39	13	20	32
57	63	49	48	91
57	67	75	48	62
57	48	42	48	32
56	46	18	48	48
55	47	81	48	05
55	58	72	20	76
55	31	66	91	06
54	64	77	48	48
53	47	30	48	59
53	45	63	08	57
51	57	35	48	86
50	52	41	20	89
49	42	72	01	57
49	42	53	48	72
49	51	41	20	38
48	57	38	03	48
48	55	82	03	96
48	34	40	78	48
47	52	47	48	67
47	42	35	48	16
46	46	47	48	48
46	38	26	78	01
45	30	55	48	16
45	30	21	20	48
44	44	54	48	05
43	84	65	20	16
43	47	21	48	10
43	45	71	48	59
43	82	69	20	62
43	54	76	48	16
42	49	36	20	23
42	47	50	78	32
42	38	72	08	48
41	41	26	48	70
41	41	13	48	32
40	51	39	78	23
39	37	58	20	08
39	50	37	48	62
39	31	17	48	16
38	35	54	48	59
37	52	37	48	16
37	28	34	48	72
36	39	43	48	29
36	54	70	48	82
34	34	24	20	82
34	43	52	48	48
33	24	48	78	39
33	79	73	20	06
32	38	29	48	16
31	23	28	78	16
31	34	20	48	16
31	28	33	78	48
30	35	46	48	10
29	35	10	78	48
29	25	16	48	48
29	32	05	78	16
28	40	09	48	32
28	26	25	48	32
27	26	14	48	29

PERCENTILE RANKS

1	2	3	4	5
27	34	20	48	23
27	22	31	91	02
27	32	65	48	67
25	25	14	78	32
25	32	22	48	25
24	42	32	48	48
23	35	44	48	32
23	19	29	48	48
23	33	60	20	16
22	22	22	78	62
22	30	27	48	32
22	25	08	20	72
20	37	33	78	32
20	21	43	48	35
20	21	19	78	05
20	22	49	78	16
19	31	29	48	48
19	16	36	48	84
18	11	48	78	23
18	25	23	48	32
17	29	43	78	48
17	17	17	78	48
17	19	57	48	16
17	19	40	48	67
17	27	59	78	95
15	15	30	48	08
15	19	10	48	32
15	27	23	20	89
15	15	18	20	28
14	15	09	78	72
13	10	05	78	08
13	11	08	48	48
13	14	28	78	04
12	15	15	78	03
12	11	07	78	35
12	07	19	95	10
11	16	11	48	16
09	11	06	78	87
09	04	02	01	48
09	06	12	48	39
09	12	04	78	23
08	09	04	78	48
08	10	25	48	05
08	08	06	78	48
07	11	02	78	23
06	06	03	48	48
06	04	14	78	28
05	05	31	48	16
05	06	07	48	48
04	09	39	48	25
04	05	32	78	38
03	03	42	78	48
03	02	12	20	23
02	03	15	95	23
02	02	03	20	48
01	01	01	48	48
01	01	01	08	25
01	01	01	78	48

Table 3

Intercorrelations of Five Over-all Performance Measures

		Judged Weights (1)	Equal Weights (2)	Statistical Weights (3)	Overall Rating of Criteria Folders (4)	Overall Rating of Physicians by Name (5)
Judged Weights	(1)	--				
Equal Weights	(2)	98*	--			
Statistical Weights	(3)	90*	86*	--		
Overall Rating of Criteria Folders	(4)	51*	55*	46*	--	
Overall Rating of Physicians by Name	(5)	46*	48*	47*	45*	--

Note: N = 205
* P < .01

three scores which represented different methods of criterion score weighting--judged, equal, and statistical weights (variables 1, 2, and 3 in Table 3)--were highly related to one another. Variable 4, the rating of the coded, anonymous criterion folders, and variable 5, the independent appraisal of each physician by name, demonstrated somewhat lower relationships with the first three combined criterion scores and with each other. Nevertheless, these correlations were still significant far beyond the .01 level. It is challenging to note that the judges who looked through the criterion folders to make their judgments must have given quite "different subjective weights" to many of the criterion scores that were used in the first three explicit methods for weighting and combining these 80 types of criterion data.

The ratings of specialists by name obtained from the five physician judges, were necessarily subjective and global in nature. They did, however, provide an estimate of how closely the criterion summary scores, based on 80 previously collected criteria, reflected actual performance. Our results indicate that the composite criterion scores hold more promise than many critics have thought as reasonably accurate estimates of on-the-job performance, for use until they can be further extended and refined. If only the first three methods had been used, the weighting problem might have been seen as a fairly trivial problem, but inclusions of the last two overall subjective rating methods brought out some important differences in computing weighting of the 80 separate criteria in the search for consistency on a composite overall rating.

An Examination of Multiple Approaches to Weighting the Performance Criteria

Five methods of ranking physicians have already been discussed. Three of those methods (judged, equal, and statistical weights) were the results of weighting each criterion variable according to a specific weighting system. However, these three sets of weights represent only a few of the ways in which the criteria could have been weighted. The purpose of the next phase of the present study was to explore additional ways of assigning weights to the criteria.

The Correlational Approach to Weighting

With the data available, it was possible to generate five sets of "correlational" weights. This procedure simply involved correlating each of the 80 criterion measures with one of the five overall summary scores mentioned in the first part of this paper. Thus one set of correlational weights was obtained by correlating each criterion with the judged weight summary score, the higher the correlation, the greater the weight. In the same manner, a set of correlational weights was calculated by correlating each criterion with the other four summary scores, namely, the equally weighted summary score, the statistically weighted summary score, the card ratings of specialists, and the criterion folder rating.

Other Weighting Approaches

Two other methods of weighting have already been discussed in an earlier section of this paper. The judged weights were used in determining the first summary score used to rank physicians in Table 2. The statistical approach i.e., weighting by multiple regression, was also used in weighting the 80 criteria, and was obtained by a regression of the 80 criteria against

the criterion folder score. The last set of statistical weights was calculated by multiple regression techniques, this time against the overall rating of physicians by name.

Thus, the eight different sets of weights listed below were obtained (equal weights were not used in this analysis since every criterion measure received a weight of 1 in this method).

Eight Sets of Weights

General Method of Weighting	Weight Set #	Specific Source
I CORRELATIONAL	1	Correlation between each criterion and the equal weight summary score.
	2	Correlation between each criterion and the judged weight summary score.
	3	Correlation between each criterion and the best-statistical summary score.
	4	Correlation between each criterion and the rating of the criterion folder.
	5	Correlation between each criterion and the rating of each specialist by name.
II JUDGMENT	6	Weight assigned by expert judgment of the criteria in terms of importance.
III BEST STATISTICAL	7	Beta weights from regression of 80 criteria on the criterion folder rating.
	8	Beta weights from regression of 80 criteria on the rating of specialists by name.

Tables 4 and 5 present an analysis of the 80 criteria in terms of the eight sets of weights. For clarity of presentation and as a better indication of overall standings of the criteria, the eight sets of weights were summed to determine a grand total importance level for each criterion. In both Tables 4 and 5, the 80 criteria are divided into three levels of importance.

It was also of interest to note how much variability there was in the weights assigned to each criterion. This provided an indication of the stability (consistency) or the instability of the performance measures. Table 4 presents the criteria which were fairly stable and consistent. These measures tended to receive the same weights regardless of which of the 8 methods was used to assign weights. Table 5 presents criteria found to be rather unstable or inconsistent, receiving comparatively different weights across the 8 methods.

An examination of Tables 4 and 5 provides some idea of the present status of our criterion analysis. In Table 4 some of our 80 criterion measures are both quite stable and important, while others tend to be stable and unimportant, not really worth a great deal of further consideration. These two extreme instances are fairly clear cut. In Table 5, however, many of our criteria are less clearly delineated. That is, they may be given considerable weight or importance by some weighting methods, but little by others. These discrepancies may give us some insight into the weights themselves, but also provide some real problems in deciding which criteria in the top two levels of importance are worthy of further consideration and refinement.

TABLE 4

Criteria Receiving Highly Similar Weights by Eight Methods of Weight Assignment
(Criteria are Arranged in Rank Order of Importance)

CONSISTENT CRITERIA--OF GREAT IMPORTANCE	Judg	Obj*
Number of times nominated as outstanding contributor by colleagues	x	
Number of times nominated as preferred consultant by colleagues	x	
Number of times nominated as preferred consultant by General Practitioners	x	
Judged quality of national board certification	x	
Total number of listings in honorary compendiums		x
Self-reported number of contributions made to medicine	x	
Number of formal responsibilities physician has in hospitals		x
Expert panel rating of overall performance based on all available information	x	
Grade-point average for first two years of medical school		x
CONSISTENT CRITERIA--OF MODERATE IMPORTANCE		
Grade-point average for last two years of medical school	x	
Number of current memberships in civic and political organizations		x
Average number of house calls made per week		x
Interviewer rating of likeability	x	
Years of experience since receiving M.D.		x
Number of hospitals in which physician maintains courtesy privileges	x	
Average number of speeches on medical topics to laymen groups per year		x
Interviewer rating of physician's involvement in this project	x	
Extent of physician's experimental use of drugs provided by drug detail men		x
Proportion of office patients that fail to pay physician for services rendered		x
CONSISTENT CRITERIA--OF LITTLE IMPORTANCE		
Mobility rate in professional positions since receiving M.D.		x
Self-reported number of non-medical contributions to society	x	
Self-estimated average socioeconomic level of patients		x
Self-estimated value of office equipment		x
Number of M.D. assistants on physician's ancillary staff		x
Number of technicians on physician's ancillary staff		x
Characteristics considered important for success: number of "uncommon" responses	x	
Number of patients seen per day		x
Proportion of office patients treated without charge		x
Judged quality of hospitals in which physician maintains courtesy privileges	x	
Characteristics considered important for success: number of "common" responses	x	
Characteristics considered important for success: "commonness" of responses	x	
Self-rating of success in medicine	x	x
Percentage of patients on which consultations are requested		x
Physician's evaluation of usefulness of drug detail men		x
Number of hospitals in which physician works		x
Number of nurses on physician's ancillary staff		x
Extent to which physician plans and maintains leisure time activities		x
Individual practice rather than group, clinic or hospital practice		x
Number of different residency hospitals		x
Number of clerical, administration, and janitorial workers on staff		x
Age at which received M.D.		x
Number of years between receiving M.D. and receiving national board certification		x
Hospital practice rather than individual, group, or clinic practice		x

*Judg = Criterion dependent primarily upon judgment
Obj = Criterion gathered from primarily "objective" sources

TABLE 5

Criteria Weighted Differently by Each of the Eight Weighting Methods
(Criteria Are Arranged in Rank Order of Importance)

UNSTABLE CRITERIA--OF GREAT IMPORTANCE	Judg.	Obj.*
Number of times nominated as outstanding contributor by college of medicine faculty	x	
Present college of medicine clinical faculty rank		x
Average number of journal publications per year since receiving M.D.		x
Average judged quality of societies in which membership is current	x	
Average judged quality of scientific and professional awards received during career	x	
Total number of papers presented at scientific and prof. meetings during career		x
Rating of "clinical excellence" by medical college department head	x	
Number of scientific and professional journals reviewed regularly		x
Number of times during career invited to serve as editor of scientific journal		x
Average judged quality of self-reported contributions made to medicine	x	
Times during career invited to serve as editor of scientific journal		x
Number of subscriptions to scientific and professional journals		x
Average number of formal medical consultations called into monthly		x
Contribution to publications as indicated by sr. vs. jr. author status		x
Number of research projects with which involved during career		x
Interviewer rating of physician's involvement in this project	x	
UNSTABLE CRITERIA--OF MODERATE IMPORTANCE		
Gross income from medical profession		x
Number of articles in professional journals read in detail each month		x
Average judged quality of hospital responsibilities	x	
Average amount of times spent with patients on first visit		x
Average number of hospitalized patients		x
Number of years spent in residency		x
Unusual techniques of keep abreast		x
Average number of informal medical consultations called into monthly		x
Average judged quality of hospitals in which physician works	x	
Number of current memberships in social and avocational organization		x
Undergraduate grade point average	x	
Average amount of time spent in explaining diagnoses to patients		x
Average amount of vacation taken annually		x
Average number of society meetings attended annually		x
Number of postgraduate courses taken during career		x
UNSTABLE CRITERIA--OF LITTLE IMPORTANCE		
Overall occupational satisfaction	x	
Number of refresher courses tak during career		x
Degree to which physician adheres to patient appointment schedule		x
Degree to which physician considers psychological factors in diagnoses		x
Number of current memberships in scientific and professional societies		x
Average number of hours per week devoted to medical practice		x

*Judg. = criterion dependent primarily upon judgment
Obj. = criterion gathered from primarily "objective" sources

Summary and Conclusions

The first phase of the present study examined three methods of synthesizing and combining 80 criteria of physician performance into single composite scores. In addition, an independent rating of the performance of each of the 205 specialists in the sample was obtained as was a rating of a coded anonymous criterion folder. Thus, five overall scores were obtained for each physician. These scores were then arranged in rank order in profile form for each specialist to indicate the relative standing of each subject on all five scores. The three overall scores based on the differential weightings of the 80 performance measures were highly correlated with each other, and moderately correlated with the other two independently obtained overall scores.

The second phase of the study dealt with how the 80 criteria mentioned above were weighted by three distinct techniques including judgmental, correlational, and best-statistical methods of weight assignment. The criteria were subsequently classified in terms of their importance and variability under all weighting methods.

The major conclusions of the study were: (1) The 80 performance criteria yielded remarkably similar summary scores when summed and weighted by three different techniques. (2) Substantial relationships were demonstrated between the summary scores based on the criteria per se and the independent appraisals of physician performance. (3) Many of the criteria examined in this study were extremely variable over different weighting methods while others were quite stable, an indication of the complexity of the whole weighting problem and the entire problem of getting one best overall score.