DOCUMENT RESUME

ED 246 070 TM 840 341

AUTHOR Sigmon, Gary L.; Halpin, Gerald
TITLE Application of Judgmental Standard Setting Procedures to Vocational Evaluation Competency Statements by Rehabilitation Field Personnel and Educators.
PUB DATE 26 Apr 84
NOTE 26p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New Orleans, LA, April 24-26, 1984).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Certification; *College Faculty; *Cutting Scores; Evaluation Methods; *Evaluators; *Minimum Competencies; *Vocational Evaluation; Vocational Rehabilitation
IDENTIFIERS *Angoff Methods; *Ebel Method; Standard Setting

ABSTRACT

Traditionally, judgmental standard setting methods have been used exclusively at the test item level. In this study, the Ebel and Angoff methods of standard setting were utilized to determine minimum competency standards on a list of 175 identified competency statements for vocational evaluators. The following research questions were addressed: (1) Do the Ebel and Angoff methods result in the establishment of different competency standards? (2) Do vocational evaluation field personnel and vocational evaluation educators establish different standards? and (3) Are the standards determined by the different methods moderated by group membership? The raters were 19 vocational evaluation field personnel and 24 vocational evaluation educators from university settings. Significant differences were found between rating methods across groups; the Ebel method resulted in a higher mean competency standard. The groups of raters also set significantly different competency standards across methods; the field personnel established a higher mean competency standard. The importance of this group difference was minimized due to the significant interaction between group and method; here, the group difference was greater using the Angoff method. These results substantiate claims that these standard setting methods may be applied to measures other than test items. (BW)

Application of Judgmental Standard Setting Procedures to Vocational Evaluation Competency Statements by Rehabilitation Field Personnel and Educators

Gary L. Sigmon

Appalachian State University

Gerald Halpin

Auburn University

Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, Louisiana, April 26, 1984.

Application of Judgmental Standard Setting

Procedures to Vocational Evaluation Competency

Statements by Rehabilitation Field Personnel and Educators

Methods of standard setting have been described and categorized by numerous authors in the fields of educational and psychological measurement (Buck, 1977; Hambleton, Powell, & Eignor, 1979; Livingston & Zieky, 1978; Meskauskas, 1976; Popham, 1978). Prevalent in the literature are three general classifications of methods including normative, criterion, and judgmental methods (Buck, 1977; Hambleton, Powell, & Eignor, 1979). Normative procedures produce standards that have been established as a result of comparison with a norm group, i.e., test scores of individuals that have already been assessed. The resulting standard then may be related to a particular point on the test score scale, such as one standard deviation below the mean, or at the sixteenth percentile (Glass, 1978). Criterion based standards are those which..."are built upon the existence of an outside criterion measure, performance measure, or true ability distribution" (Hambleton, Powell, & Eignor, 1979, p. 43). Judgmental methods may be employed without the use of examinee data to set a standard. Unlike normative and criterion methods, decisions are made relative to test item content with judgmental procedures where raters have access to test items and not external criterion such as examinee data. The most prevalent judgmental procedures for setting

standards are those described by Ebel (1972), Angoff (1971), and Nedelsky (1954).

Not unlike other occupations requiring the certification or other documentation of minimal competence, the field of vocational evaluation is presently developing procedures for the formal certification of vocational evaluation personnel (Coffey & Mason, 1980). Actually, the certification of vocational evaluation personnel has been an issue for a number of years (Baker, 1979; Brubaker, 1979; Coffey, 1978; Coker, 1980; McDaniel, 1978; Stout, 1969). This movement within the field of vocational evaluation led by the Vocational Evaluation and Work adjustment Association (VEWAA), a division of the National Rehabilitation Association, began with efforts to determine professional standards for vocational evaluators. VEWAA has been involved with standard development for at least 10 years (Couch, 1971).

A prerequisite to the certification issue and an area of much concern involved the establishment of certification criteria. A favored approach among the professionals in the field of vocational evaluation has been to determine what competencies are necessary to function effectively as a vocational evaluator. For example: Coffey (1978) identified more than 2500 competency statements that related to the work of vocational evaluators from educators, practitioners of vocational evaluation, and literature; Sink, Porter, Rubin, and Painter (1979) examined differences and

similarities of competency requirements for vocational evaluators and rehabilitation counselors; and the identification and classification of competencies was considered critical to the process of competency-based certification by Gannaway and Sink (1979).

From the massive list of over 2000 vocational evaluator competency statements compiled by Coffey (1978), a final list of 175 competencies was extracted through combining, categorizing, and omitting statements (Coffey, 1978). As the profession of vocational evaluation and its proponents for certification have moved in the direction of formalizing certification procedures, a critical need emerged to apply a systematic procedure to this list in order to refine and accurately address the most relevant or core competencies required for the profession. With the decision by the Vocational Evaluation and Work Adjustment Association (VEWAA) to proceed with developing certification procedures for vocational evaluation and work adjustment personnel, came the necessity to refine and prioritize competency requirements for the assessment of minimal competency to engage in the work of these professions. This study was an attempt to provide additional information relative to core competency requirements as perceived by rehabilitation evaluation educators and field personnel representing the practicing professionals in vocational evaluation. The problem addressed in this study was the need to refine and prioritize competency areas

within the field of vocational evaluation for the purpose of aiding in the development of vocational evaluation professional certification procedures.

Objectives of the Inquiry

Traditionally, judgmental standard setting methods have been used exclusively at the test item level. In this study the Ebel and Angoff methods of standard setting were utilized to determine minimum competency standards on a list of 175 identified competency statements for vocational evaluators. The following research questions were addressed:

1. Do the Ebel and Angoff methods of standard setting result in the establishment of different competency standards when applied to competency statements for vocational evaluators?

2. Do vocational evaluation field personnel and vocational evaluation educators establish different competency standards when utilizing these methods?

3. Are the competency standards determined by the different methods moderated by group membership?

Method

Subjects were 43 raters representing two professional groups. The first group consisted of 19 vocational evaluation field personnel, and the second group was composed of 24 vocational evaluation educators from university settings. The practitioners were state chapter presidents of the Vocational Evaluation and Work

Adjustment Association, and the educators were identified from the current membership directory of the National Council on Rehabilitation Education.

Each rater in each group rated each of the 175 competency statements utilizing first the Ebel method followed by the Angoff method. The raters worked independently as the instructions, personal data sheets, rating forms, and competency statement lists were distributed via mail.

For the Ebel method nine categories were listed on the rating form, and the raters were instructed to check one of the nine categories indicating whether the competency statement being rated was considered essential, important-easy, important-medium, acceptable-easy, acceptable-medium, acceptable-hard, questionable-easy, questionable-medium, or questionable-hard. These categories made it necessary for the raters to decide upon the relevance of the competency as well as the difficulty level of the competency statement. These decisions were to be made in reference to beginning vocational evaluators minimally competent in the profession.

The second method, Angoff, required raters to estimate the percentage of beginning vocational evaluators, minimally competent in the profession, who in their opinion would be able to demonstrate the listed competency statement.

The Ebel-Angoff ordering of the ratings was prescribed in an attempt to minimize contamination of the ratings due to carry-over effects. The Ebel procedure required the rater to make general nonnumerical ratings, whereas the Angoff procedure mandated that a specific number, (a proportion) be assigned to the competency statement. Both methods are concerned with competency difficulty, but in the Ebel method it is estimated indirectly and is less obvious to the raters than in the Angoff procedure.

After all competency statements were rated by the two procedures, the raters opened a separate envelope with instructions to provide a percentage representing an expected success rate for each of the nine Ebel categories. In essence the raters were asked

---

Insert Table 1 about here

---

to determine what percentage of the competency statements allocated to each of the nine categories should be required to determine minimal competency.

To determine if the use of different methods resulted in different competency standards across goups or if different groups determined different competency standards with the same method, a 2 x 2 mixed design analysis of variance was performed with group membership representing the between group factor and rating method

the within group factor. The dependent variable was the minimum competency standard obtained from each rater with each method.

Results and Conclusions

Insert Table 2 about here

Significant differences, $F(1, 41) = 58.62$, $p < .001$, were found between rating methods across groups. Use of the Ebel method resulted in a mean competency standard (120.75) significantly higher than that set with the Angoff method (84.09).

Insert Table 3 about here

The groups of raters also set significantly different, $F(1, 41) = 4.84$, $p < .05$, competency standards across methods. The mean competency standard established by the field personnel (100.88) was significantly higher than the mean competency standard set by the educators (97.31). The importance of this determined group difference was minimized due to the significant interaction between rater group and rating method, $F(1, 41) = 5.18$, $p < .05$. Here, the

---

Insert Figure 1 about here

---

difference between the mean competency standards set by the practitioner group and the educator group was greater for the

Angoff method (96.43 - 74.33 = 22.10) than for the Ebel method (121.34 - 120.29 = 1.05).

Additionally, interrater reliability was determined for the Ebel (r = 97) and Angoff (r = .93) methods along with the intercorrelation between the methods (r = .83). The average

Insert Table 4 about here

competency ratings for the Ebel and Angoff methods were correlated with the average ratings of importance of these competencies obtained in a previous study by Coffey. The Ebel and Angoff

Insert Table 5 about here

competency ratings correlated .88 and .70, respectively, with the importance ratings obtained by Coffey.

Educational Importance of the Study

Historically, no precedent has been set in the utilization of judgmental standard setting procedures with any data other than test items. Various studies have been conducted demonstrating the use of these methods with competency examinations (Andrew & Hecht, 1976; Archambault, Behuniak, & Gable, Note 2; Behuniak, Gable, & Archambault, Note 3; Colton & Hecht, Note 4; Koffler, 1980; Poggio, Glasnapp, & Eros, Note 5; Sigmon, Halpin, Halpin, & Sides, Note 6; Skakun & Kling, 1980). In the present study the judgmental methods

developed and described by Ebel (1972) and Angoff (1971) were utilized with competency statements which represents an entirely new and innovative application of these standard setting methods. Personal communication with the developer of one of these techniques (Ebel, Note 7) reinforced the initiation of this study. Other justification for this approach was provided by Stewart and Maslow (Note 1) in the following:

> Procedures other than written tests do not provide such convenient units as 'test items' as a basis for standards setting.... Nevertheless, it is possible to apply to other measures the same concern to standardization, reliability, and control of the judgmental process. (p. 15)

In this unique application of judgmental standard setting procedures to competency statements, results obtained in the present study were similar to the results found with the traditional application of these methods to test items. As in previously cited studies (e.g., Andrew & Hecht, 1976; Skakun & Kling, 1980; Colton & Hecht, Note 4) where method differences were found among various judgmental standard setting procedures, this study resulted in the formulation of the conclusion that the Ebel and Angoff methods may be utilized to determine competency standards, but will result in different competency standards established.

The results of this investigation tend to substantiate the statements of Ebel (Note 7) and Stewart and Maslow (Note 1) that these procedures (standard setting) may be applied to measures other than test items. The significant findings in the reliability and validity studies within the present investigation also support the investigators' contention that item judgmental standard setting procedures (i.e., Ebel and Angoff) can be suc essfully applied to competency statements as well as test items.

The demonstrated applicability of item judgmental standard setting procedures, i.e., Ebel and Angoff, to measures other than test items provided new insights into the versatility of these methods. Through the use of these procedures at the competency level, the task of developing tests for competency certification became better defined. As test items are developed for each competency identified, the decisions to be made relative to overall competency for a specific profession may be made at a level that is substantially more specific to the profession than at the test item level.

The analyses performed and results obtained in the present investigation of vocational evaluator competency statements may be utilized in the establishment of either noncompensatory or compensatory certification models (Allen & Yen, 1979). Noncompensatory certification refers to requiring a minimal competency level for each of an established number of specific

subareas within a profession. In a compensatory model, only a total performance level or mastery level is required without requiring minimally expected levels within subareas. For example, if a university graduate program requires a total Graduate Record Examination (GRE) score of 1000 for admission without requiring specific scores for the verbal or mathmatics sections, the compensatory model is being utilized. However, if the school requires a GRE Score of 1000 with minimum scores of 450 on each area a noncompensatory model is being applied (Allen & Yen, 1979).

Theoretically, the methodology utilized in this investigation can be utilized in other areas to increase, possibly insure, the presence of content validity in future test development for professional certification, licensure, or registry.

The application of the results of this study to a certification procedure for purposes of standard setting precipitates the decision of which competency standards to utilize. Consistently, the different methods, Ebel and Angoff, produced different competency standards. A conservative approach might be to choose the results produced by the Angoff procedure averaged across the two groups of raters, since the Angoff procedure consistently resulted in lower competency cut-offs. However, the educator group set the lowest standards utilizing the Angoff method consistently. Another approach may be to select the results produced by the Ebel procedure averaged across rater groups. In this study, as in others, (Andrew

& Hecht, 1976; Sigmon, Halpin, Halpin, & Sides, Note 6) the Ebel procedure consistently resulted in higher cut-offs, but was also the most stable across rater groups. Other researchers have suggested combining methods or utilizing several methods, to arrive at the cut-off decision (Koffler, 1980; Livingston & Zieky, 1978). By combining the two methods and averaging across rater groups the resulting competency standards are based upon decisions made by both groups of raters utilizing both methods. For example, in this study the field personnel established competency standards of 96 competency statements and 121 competency statements via the Angoff and Ebel methods, respectively. The average across the two methods is 108 for the field personnel group. Similarly, for the educator group competency standards via the Angoff and Ebel procedures were 74 and 120, respectively, for an average across methods of 97. Those two averages could then be combined and averaged for a competency standard of 102 of 175 competency statements, which results in a cut-off of 58% (Sigmon, Halpin, Halpin, & Sides, Note 6). This figure could then be utilized for a compensatory model for certification procedures.

Reference Notes

1. Stewart, J. T., & Maslow, A. P. (April, 1981). The setting for standard setting and the standard for standard setting. Paper presented as part of a symposium entitled Standard setting: procedures, practices, and contraints, at the annual meeting of the National Council on Measurement in Education, Los Angeles.
2. Archambault, F., Behuniak, P., & Gable, R. (April, 1981). A comparison of the Angoff and Nedelsky procedures for setting standards on reading and mathematics proficiency examinations. Paper presented at the annual meeting of the National Council on Measurement in Education, Los Angeles.
3. Behuniak, P., Gable, R., & Archambault, F. (April, 1981). A comparison of dichotomized and trichotomized proficiency test scores. Paper presented at the annual meeting of the National Council on Measurement in Education, Los Angeles.
4. Colton, D. A., & Hecht, J. T. (April, 1981). A preliminary report on a study of three techniques for setting minimum passing scores. Paper presented at the annual meeting of the National Council on Measurement in Education, Los Angeles.
5. Poggio, J. P., Glasnapp, P. R., & Eros, D. S. (April, 1981). An empirical investigation of the Angoff, Ebel and Nedelsky standard setting methods. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles.

6. Sigmon, G. L., Halpin, G., Halpin, G., & Sides, E. (April, 1981). Use of judgmental procedures by groups of raters to set minimum competency standards. Paper presented at the annual meeting of the National Council on Measurement in Education, Los Angeles.

7. Ebel, R. (March 4, 1981). Personal communication.

References

Allen, M., & Yen, W. (1979). Introduction to measurement theory. Monterey, CA: Brooks/Cole.

Andrew, B. J., & Hecht, J. T. (1976). A preliminary investigation of two procedures for setting examination standards. Educational and Psychological Measurement, 36, 35-50.

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In D. L. Thorndike (Ed.), Educational measurement. Washington, DC: American Council on Education.

Baker, R. J. (1979). A dialogue on certification of vocational evaluators and adjustment services specialists: A response from a VE and WA educator. Vocational Evaluation and Work Adjustment Bulletin, 12(1), 4-6.

Behuniak, P., Archambault, F., & Gable, R. (1982). Angoff and Nedelsky standard setting procedures: Implications for the validity of proficiency test score interpretation. Educational and Psychological Measurement, 42, 247-255.

Brubaker, D. (1979). A dialogue on certification of vocational evaluators and adjustment services (work adjustment) specialists: The appraisal of one rehabilitation counselor leader. Vocational Evaluation and Work Adjustment Bulletin, 12(1), 2-3.

Buck, L. S. (1977). Guide to the setting of appropriate cutting scores for written tests: A summary of the concerns and procedures. Washington, DC: Personnel Research and Development Center, United States Civil Service Commision.

Coffey, D. (1978). Vocational evaluator competencies and their relevance as perceived by practitioners and educators in vocational evaluation. Unpublished doctoral dissertation, Auburn University.

Coffey, D., & Mason, A. (1980). Certification of vocational evaluators and work adjustment specialists: Its history and its future. Vocational Evaluation and Work Adjustment Bulletin, 13(3), 83-85.

Coker, C. C. (1980). Certification of vocational evaluators: results of the survey. Vocational Evaluation and Work Adjustment Bulletin, 13(3), 92-97.

Couch, R. H. (1971). The involvement of VEWAA in training, standards, and research. Vocational Evaluation and Work Adjustment Bulletin, 4(2), 26-29.

Ebel, R. L. (1972). Essentials of educational measurement. Englewood Cliffs, NJ: Prentice-Hall.

Gannaway, T. W., & Sink, J. M. (1979). An analysis of competencies for counselors and evaluators. Vocational Evaluation and Work Adjustment Bulletin, 12(3), 3-15.

Glass, G. V. (1978). Standards and criterion. Journal of Educational Measurement, 15, 237-261.

Hambleton, R. K., Powell, S., & Eignor, D. R. (1979). Issues and methods for standard setting. In Laboratory of Psychometric and Evaluative Research Report No. 70. Amherst, MA: School of Education, University of Massachusetts.

Koffler, S. L. (1980). A comparison of approaches for setting proficiency standards. Journal of Educational Measurement, 17, 167-178.

Livingston, S. A., & Zieky, M. J. (1978). Manual for setting standards on the basic skills assessment tests. Princeton, NJ: Educational Testing Service.

McDaniel, R. S. (1978). Vocational evaluation: On becoming a profession. Vocational Evaluation and Work Adjustment Bulletin, 11(3), 29-38.

Meskaushas, J. A. (1976). Evaluation models for criterion-referenced testing: View regarding mastery and standard setting. Review of Educational Research, 46, 133-158.

Nedelsky, L. (1954). Absolute grading standards for objective test. Educational and Psychological Measurement, 14, 3-19.

Popham, W. J. (1978). Setting performance standards. Los Angeles, CA: Instructional Objectives Exchange.

Sink, J. M., Porter, T. L., Rubin, S. E., & Painter, L. C. (1979). Competencies related to the work of the rehabilitation counselor and vocational evaluator, Vol. 1. The University of Georgia, Athens.

Skakun, E. N., & Kling, S. (1980). Comparability of methods for setting standards. Journal of Educational Measurement, 17, 229-235.

Stout, J. (1969). Certifying the evaluator. Vocational Evaluation and Work Adjustment Bulletin, 2(4), 4-6.

Table 1

Expected Success Rates by Field Personnel and Educators

Item Category	Expected Success Rate-Field Personnel	Expected Success Rate-Educators
Essential	.91	.87
Important-Easy	.87	.83
Important-Medium	.76	.82
Acceptable-Easy	.73	.67
Acceptable-Medium	.65	.63
Acceptable-Hard	.57	.56
Questionable-Easy	.30	.31
Questionable-Medium	.22	.25
Questionable-Hard	.16	.16

Table 2

Analysis of Variance for Rater Groups and Methods

Source	df	MS	F
Rater groups	1	2841.55	4.84*
Error (between group)	41	586.74	-
Methods	1	26629.55	58.62***
Rater x Method Interaction	1	2351.09	5.18*
Error (within group)	41	454.29	-

*$p < .05$. ***$p < .001$

Table 3

Means by Groups and Methods

Group	Method		Row
	Angoff	Ebel	Means
Field Personnel	96.43	121.34	100.88
Educators	74.33	120.29	97.31
Column Means	84.09	120.75	

*n = 175 competency statements

Table 4

Interrater Reliability for Each Group by Each Method

	Ebel	Angoff
Field Personnel	.92	.84
Educators	.95	.89
Combined Groups	.97	.93

Table 5

Sigmon-Coffey Data Relationships

Sigmon Group/Method[1]	Coffey Group[2]	Rho*
Combined/Angoff	Combined	.70
Combined/Ebel	Combined	.88
Field Personnel/Angoff	Practitioners	.63
Educators/Angoff	Educators	.63
Field Personnel/Ebel	Practitioners	.81
Educators/Ebel	Educators	.87
Field Personnel/Angoff	Educators	.63
Educators/Angoff	Practitioners	.67
Field Personnel/Ebel	Educators	.79
Educators/Ebel	Practitioners	.85

*All correlations were significant at $p < .001$

[1]Sigmon grouping = combined (19 Field Personnel and 24 Educators)

[2]Coffey grouping = combined (96 Practitioners and 20 Educators)

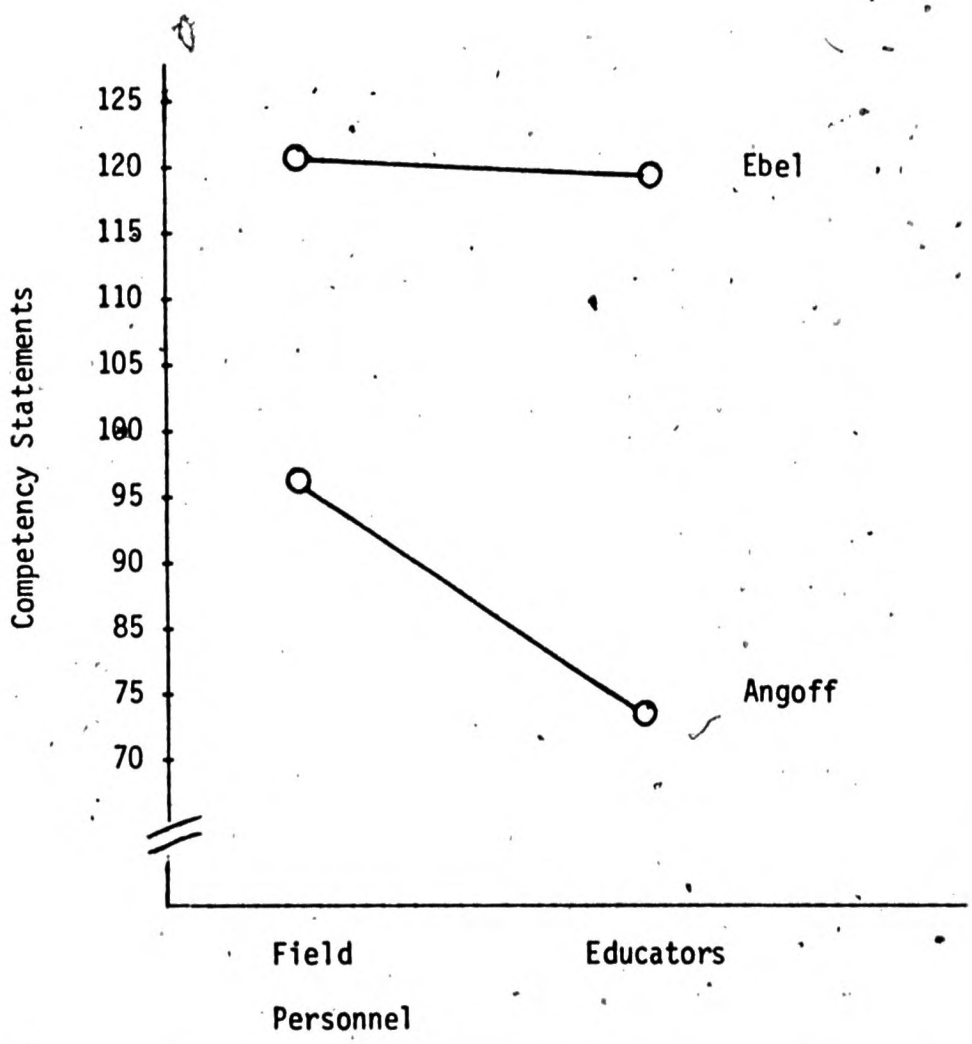

Figure 1. Mean competency standard for groups by method.