DOCUMENT RESUME

ED 142 579 TM 006 403

AUTHOR Carey, Lou M.
TITLE An Investigation of the Validity of Using Self-Evaluation Instruments to Identify Instructional Needs.
NOTE 11p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Comparative Analysis; Elementary Secondary Education; Inservice Teacher Education; Needs Assessment; Performance; Performance Based Teacher Education; *Performance Tests; Predictive Validity; *Self Evaluation; Teachers; *Teaching Skills; Testing Problems; Test Results; *Test Validity

ABSTRACT

The validity of interpreting teachers' perceptions of their performance as an index of their actual performance was examined. Two matching instruments were constructed; each contained 72 items in 6 categories of skills. One assessed teachers' perceptions of their competence on behaviorally stated generic teaching skills, and the other assessed teachers' actual performance on the same skills. The items represented verbal information, concept identification, or problem-solving skills that teachers need to perform the skill objectives; and skills that could realistically be assessed using pencil and paper questions. One hundred seventy-five classroom teachers were paid to participate in the study. Results showed that teachers' perception scores were significantly higher than their actual performance scores in all six content areas. Using teachers' perception scores to predict actual performance scores on teaching skills appeared to be an invalid practice; this finding held true for three different types of questions: recall of verbal information, concept identification, and problem-solving questions. It was also recommended that instructional needs of teacher education programs be determined by the teachers' actual performance rather than their perceived skills. (MV)

AN INVESTIGATION OF THE VALIDITY OF USING SELF-EVALUATION INSTRUMENTS TO IDENTIFY INSTRUCTIONAL NEEDS

Lou M. Carey

Florida State University

Background Information

The performance-based or competency-based teacher education movement is becoming instrumental in the redefinition of goals for inservice and preservice teacher training institutions. More institutions are moving toward a belief that they should provide specific skills and knowledge which classroom teachers will need to effectively perform in today's schools. The redefinition of goals as performance skills for classroom teachers has enabled teacher training institutions to evaluate current curriculum and instruction and begin to redefine both in terms of teachers' needs. This new emphasis on performance skills for classroom teachers is also being felt in teacher centers.

Teacher centers are now being established by many states and local districts to be responsible for the inservice education of classroom teachers. Many are operated by local school districts while some are sponsored jointly by school districts, universities, and state departments of education. Like many new organizations, teacher centers are undergoing the processes of establishment, definition, and maintenance. The information needs for such early planning have been so great that in many districts special funding has been allocated for needs assessment studies to help gather this information. Legislation has been passed in many states to establish teacher education centers, and funds have been provided to enable needs assessments for program definition in teacher centers. In Florida, for example, special needs assessment programs are either planned, currently underway, or recently completed.

One purpose for these needs assessment studies is to determine the skills and knowledge teachers should have and the skills and knowledge they actually do have. Skills which most teachers need but do not possess then become priority instruction for teacher centers or other agencies to provide.

Information gained through these legislated needs assessment studies should provide input for decisions which will be made about teacher training by local schools, school districts, teacher centers, and teacher education programs in colleges and universities. Because of the importance of these decisions and implications they might have, assessment procedures and instruments used to gather information during these needs assessment projects should be studied. Considerable attention needs to be given to the types of decisions that can or should be made on the basis of assessment instruments used and results obtained.

In a survey of instruments currently used in these needs assessment studies, it was discovered that most assessment instruments were designed to measure the perceptions of teachers and other school personnel about the skills and knowledge they believe teachers should possess. Several studies also use perception instruments to survey teachers' perceived performance on specified skills. Items on these perception instruments are commonly of the behaviorally stated skill statement type (Carey, 1974; Carey, 1975; Spindler, 1975; Florida State Department of Education, 1975).

Based on information about teachers' perceived instructional needs and perceived competence to perform specified skills, teacher centers, universities, and state departments of education plan to obtain instruction which will be provided to the classroom teachers in the near future.

The Problem

The purpose of this study was to examine the validity of the current practice of interpreting teachers' perceptions of their performance as an index of their actual performance on specified teaching skills. This is the procedure currently used by the state departments of education, universities, and teacher centers to assess the status of students and teachers on identified skills, to evaluate existing training programs, and to make decisions about forthcoming teacher training.

Test validation defined by Cronbach (1971) is a process to determine the accuracy of inferences made from test scores; i.e., a process to examine the soundness of interpretation of a test. Regardless of how well a test was constructed or how representative the items of a test are to the content being measured, if the test is improperly interpreted, then it is considered invalid for the purposes intended. A validity study of any particular test does not stop with developmental, administrative, or scoring procedures. It must be determined that test scores are being properly interpreted, and judgments, decisions, or predictions made from the scores are accurate.

One type of validation, the validity of use and interpretation of tests, is defined by the American Psychological Association (1966) as criterion-related or predictive validation. The procedure for this type of validation requires the comparison of test scores with a variable considered to be a direct measure of the behavior in question. Actual performance scores are used as the criterion in this study to validate the use of teachers' perceptions of their performance on specified teaching tasks.

A second method of validating the use of perception scores as accurate indicators of performance scores is to determine whether the same decisions would be made about inservice instruction needs if the decisions were based on either teachers' perceptions of

their performance or teachers' actual performance. This comparison of decisions that would be made would help determine how much confidence can be placed in inferences and decisions about classroom teachers' instructional needs which are based on measures of perception.

At this point in time when many teacher centers are attempting to respond to the needs of their professional community, it is particularly appropriate to investigate the validity of the current practice of using teachers' perceptions of their performance as indicators of skills that need to be taught in inservice training programs.

Specific hypotheses tested in the study were:

1. There are no differences between teachers' perception scores and their performance scores on behaviorally stated generic teaching skills.

2. There are no differences between teachers' perception and performance scores on verbal information questions.

3. There are no differences between teachers' perception and performance scores on concept identification questions.

4. There are no differences between teachers' perception and performance scores on problem-solving questions.

5. There is no difference in the selection of skills needed for inservice instruction when the selection is based on perception or performance scores.

Methodology

Instruments

Two matching instruments were constructed. One was to assess teachers' perceptions of their competence on behaviorally stated generic teaching skills and the other was to assess teachers' actual performance on the same generic skills.

The skills included in the perception and performance instruments were taken from instructional analyses of generic teaching skills that had been selected as important for classroom teachers to possess by over 90 percent of a group of educators representing classroom teachers, university faculty, and the Florida State Department of Education. Skills were selected from the instructional analyses for inclusion in the study if they: (a) represented either verbal information, concept identification, or problem-solving skills that teachers would need to perform the terminal objective in the skill hierarchy; and (b) represented skills that could realistically be assessed using pencil and paper questions.

The final perception and performance instruments each contained seventy-two items in six different categories of skills. These categories included skills in interacting with students, planning instruction, evaluating instruction and learning performance, communicating with

peers and parents, classroom management, and foundations of education. There were 12 questions included on each of the instruments for each of these six categories. Of the 12 questions, four required the recall of verbal information, four required concept identification, and four were problem-solving questions.

Procedures used to validate the instruments consisted of having a panel of appropriate judges determine whether: (a) the behavior stated in the skill statement on the perception test matched the behavior in the subskill in the original instructional analysis; (b) the question on the performance test measured the performance stated in the behavioral skill statement on the perception test; (c) the perceptual and performance skill items reflected skills classroom teachers should possess; (d) the questions represented skills that could be assessed using a paper and pencil instrument, and (e) the wording of the questions was clear. A small group of teachers were asked to verbally interpret the instructions for each instrument and each item on each instrument to determine whether items and instructions were clear to sample members of the target group.

Table 1 contains examples of behaviorally stated items from the perception instrument and matching test items from the performance test.

Reliability determined by coefficient alpha was .87 and .80 for the perception and the performance instruments respectively.

Sample

One hundred and seventy-five classroom teachers from Leon County, Florida participated in the study. They represented the following groups:

Age	21-25 years	26-30 years	31-40 years	41+ years	no indication
	19%	23%	20%	31%	7%

Grade Currently Teaching	Preschool-2nd	3rd-5th	6th-8th	9th-12th	no indication
	14%	23%	27%	23%	13%

University Degree	BS or BA	MS or MA	no indication
	58%	34%	8%

Procedures

Teachers were paid to participate in the study which was scheduled at three local school sites immediately following the regular school day. Teachers were asked to:

1. Complete the perception instrument by reading each behaviorally stated teaching skill statement and determine whether or not (true or false) they could perform the stated

skill.

2. Complete the performance test by actually providing the answers to each question.

3. Identify themselves only by grade level taught, years of teaching experience, and current university degree.

Nonparametric statistical procedures were used to analyze differences observed in teachers' perception and performance scores.

Results

Perception Scores and Performance Scores

Teachers' perception scores were significantly higher ($p \leq .05$) than their actual performance scores in all six content areas (Wilcoxon matched-pair signed-ranks test).

Table 2 contains mean scores and standard deviations for perception and performance scores in all six content areas.

Perception and Performance Scores Compared by Type of Question

Teachers' scores on verbal information questions, concept identification questions, and problem-solving questions on the perception instrument were significantly higher ($p \leq .05$) than their scores on matching verbal information, concept identification, and problem-solving questions on the performance test in all six content areas (Wilcoxon matched-pairs signed-ranks test).

Decisions Based on Perception Data Compared to Decisions Based on Performance Data

The six content areas were prioritized with the lowest score receiving the highest priority for inservice training. Comparing the rank order of all six content areas for perception and performance scores, the same decision about training priorities would be made one time out of six. A comparison of the rank order of the six skill areas on the perception and performance instruments is presented in Table 3.

One other comparison was made of decisions based on perception and decisions based on performance. For this comparison, all item perception scores and all item performance scores were ranked from high to low. Items with the lowest scores thus represented skills for which training priority would be highest. Ten priority skills were then selected based on perception and ten were selected based on performance. A comparison of the two priority lists indicated that six of the skills were on both lists.

When similar lists of the ten lowest priority skills were prepared, five of the skills were on both lists.

Conclusions

The first type of validity studied was criterion or predictive validity in which a set of scores is compared against a variable considered to be a direct measure of the behavior in question.

Using teachers' perception scores to predict actual performance scores on generic teaching skills appears to be an invalid practice. This finding held true for three different types of questions, namely, recall of verbal information, concept identification, and problem-solving questions.

The second type of validity considered was the comparison of the priority decisions one would make using each type of data. The prioritization of instructional needs based on perception scores was different than the prioritization based on performance scores when compared by skill area and by individual skills. Based on the dissimilarity of training decision that would have been made, it appears to be invalid to prioritize instructional needs solely on perception data.

Discussion

This study has demonstrated a need for the inclusion of performance data in the prioritization of inservice training needs. Though problems currently exist in measuring teachers' competence on specified teaching skills on a district-wide basis, the substitution of teachers' perception of their performance as a predictor for their actual performance is invalid.

Some of the problems involved in performance testing to identify priority training needs are the time involved in administering tests, the time and personnel required to hand-score teachers' free response answers, and possibly the time required to observe teachers actually perform the skills in a real or simulated setting. An investigation should be conducted into suitable procedures to select and test a representative sample of teachers in a district to determine their actual competence in selected skills.

The acquisition of accurate data concerning teachers' competence on specified teaching skills is not the only ingredient that should be used to identify inservice training priorities. For example, teachers may all fail to perform certain skills correctly, however, if no one in the district considers those skills important for successful teaching, then the fact that teachers are not competent in them becomes unimportant. The opinions of educators concerning the skills teachers *need* and the skills they most *want* included in their inservice programs are ingredients that should be considered along with performance data to prioritize skills for inservice training.

Teachers' perceptions of what programs should be provided and the value of programs

that have been provided are useful and should be sought when planning, providing, and evaluating inservice education programs. However, these opinions should not be considered a substitute for actual preinstructional, postinstructional, and follow-up assessment of teachers' *ability to perform* specified, basic teaching skills.

BIBLIOGRAPHY

American Psychological Association. *Standards for educational and psychological tests and manuals.* Washington: APA, 1966.

Carey, L. M. Leon County teachers' perceptions of generic teaching skills. Unpublished paper, Florida State University, 1975.

Carey, L. M. *A validation of self-evaluation procedures for identifying instructional needs of teacher centers.* Unpublished doctoral dissertation, Florida State University, 1976.

Cronbach, L. J. Test validation. In Thorndike, R. L. (Ed.) Educational Measurement. Washington: American Council on Education, 1971, 443-507.

Florida Department of Education, *The Florida catalogue of teacher competencies.* Chipley, Florida: Panhandle Area Educational Cooperative, 1973.

Spindler, L. J. *A comparison of two methods of determining teacher needs for inservice education.* Unpublished doctoral dissertation, Florida State University, 1976.

Table 1

Sample Matching Items from the Perception and the Performance Instruments

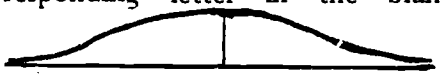
Perception Item	Performance Item
1. Given a statement of the behavior in a *verbal information learning task*, you can describe the *performance* that would prove the student could perform the task. . . . . (T) (F)	1. *Objective:* The student will learn the names of three furbearing animals. *Type of Learning:* Verbal information (labels) acceptable proof the student has learned names:
2. Given a diagram of a normal curve, you can identify the mean score. . . (T) (F)	2. Identify the *location* of the *mean* score in the illustration below, and place the corresponding letter in the blank  a b c d e f g h i j k
3. You can list different categories of information about students which you believe would be beneficial to teachers planning programs for new students . . . . . (T) (F)	3. List three different categories of information you believe would be beneficial to teachers planning programs for new students. a. _____ b. _____ c. _____

Table 2

Mean Scores and Standard Deviations for

Perception and Performance Scores in All Six Content Areas

	Perception Instrument		Performance Test	
Content Area	X	SD	X	SD
Interacting with students	7.60*	2.98	6.40	2.52
Planning instruction	8.90	2.55	5.32	2.22
Evaluating instruction and learning performance	8.71	3.19	6.31	2.93
Communication with peers and parents	8.59	2.38	2.54	2.02
Classroom management	9.73	2.41	5.09	2.49
Foundations in education	5.28	4.33	2.04	1.92

*Maximum score in any of the six content areas is 12.

Table 3

Prioritization of Instructional Needs by Skill Areas

- Based on Perception and Performance Data

Rank	Perception	$\bar{X}$	Rank	Performance	$\bar{X}$
1	Foundations	5.28	1	Foundations	2.04
2	Interacting with students	7.60	2	Communication with peers and parents	2.54
3	Communication with peers and parents	8.59	3	Classroom management	5.09
4	Evaluate instruction and performance	8.71	4	Planning instruction	5.32
5	Planning instruction	8.90	5	Evaluate instruction and performance	6.31
6	Classroom management	9.73	6	Interacting with students	6.40