INSTRUCTIONAL STRATEGIES AS PART OF THE CONTENT DOMAIN OF A CRITERION-REFERENCED TEST

Mildred A. Murray
Hillsborough County School System

INTRODUCTION

Hillsborough County Public Schools, in an effort to provide better instruction for its ECIA Chapter 1 (formerly ESEA Title I) students, is in the process of developing a criterion-referenced test designed to measure the reading behaviors essential to underachieving students. As part of this test development project, a clear delineation of an objective content base was required. Because the Chapter 1 program has been designed to supplement regular classroom activities, the objectives contained in the current basal reading system used in the County were chosen as a content base. However, the objectives proved to be either too broad or vague. The decision was made to examine the instructional processes accompanying the objectives in order to clarify them. This paper will demonstrate how the inclusion of instructional processes could be used to identify those objectives in need of refinement and/or clarification.

PERSPECTIVE

The development of criterion-referenced testing in education as documented by Hambleton and Swaminathan (1978) has spawned both argument and confusion. Perhaps the most basic of issues involved is the meaning of the word "criterion." One widely accepted definition of "criterion" is that it is related to a well defined domain of behaviors (Glaser, 1963; Popham and Husek, 1969).

Clearly, defining a behavioral domain must involve unambiguous terms, which are not arbitrary (Berk, 1980) and are relevant to the purpose of the measurement (Guion, 1977). The development of domain specifications based on a clear statement of behaviors to be measured facilitates the defining process. One finds, however, that most behavioral objectives are not clear enough for use in domain specifications (Hambleton et al., 1978). Scandura (1977) states that the basic weakness in traditional objectives is that they state a behavior, not what the student is to learn or what the teacher is to teach. Such objectives fail to measure "instructional intents" (Mager, 1973). They miss the intents because of overly generalized, vague or conflicting behaviors.

One way to improve and clarify these objectives is to use the instruction inherent in the behaviors. Such information would help define type and level of competence and modes of behavior to be observed (Scandura, 1977). The analysis of instructional strategies as part of the domain could be termed "instructional validation." This procedure, while quite logical, is usually impractical because the methods to be used to teach the behaviors are either unknown or are out of the control of the individual defining the domain. A rare opportunity for the determination of this "instructional content validity" came in the development of a criterion referenced test to be used with the Chapter 1 students in the Hillsborough County School System.

PROCEDURES

The determination of the Chapter 1 CRT content base began as a simple listing of reading objectives. The relevant objectives were those which appeared in both the Hillsborough County Reading Compendium and the current reading basal series. This group of objectives was quite large, so a prioritization of the

objectives in each basal was conducted to reduce the number of objectives.

A number of objectives to be included in the content domain were found to be vague or too broad to be useful. Terms such as "understand" or "know" were not at all uncommon. Because of these ambiguities, the instructional processes were used to further refine or clarify the objectives.

The basal series contained a large amount of material, some of which was seldom used. Processes related to that material were eliminated. The processes were ultimately defined, operationally, as the basal series itself, including the teacher direction pages and the workbooks. Further, because the posttests accompanying the series were frequently used as the culminating activities at the end of stories, they were also included.

The agreement between the instructional activities and the objectives was examined. The purpose of this activity was to identify objectives which were unclear or unrelated to the instruction. It was hoped that the broad objectives could be further focused in light of the actual intent of the instruction.

Each objective and its posttest item(s) and the instructional processes were classified, using a checklist format (Appendix A). The checklist included reading content agreement, the type of task involved, the level of cognitive processing (according to Bloom's Taxonomy) required, and the type and modality of the student response. A space for comments was also included. The completed checklists were attached to the Objective/Posttest/Instruction Agreement Sheet (Appendix B) and copies of the relevant basal and workbook pages, to form a packet for each objective at each time it was taught.

A standing committee of fifteen (15) Chapter 1 teachers working in teams of two or three, evaluated each set of materials. These teachers were selected because the Chapter 1 instructional supervisors gave them high ratings on general teaching ability and in-depth knowledge of the basal series. Two teams evaluated each set of materials independently.

The evaluations of the two teams were analyzed for discrepancies and to clarify and refine the objectives. The evaluations proved to be very valuable in determining what were the intended objectives.

RESULTS

The results of the objective/instructional process evaluations were combined and categorized as a match or non-match. While the domain definition is still continuing, some very important and surprising findings were obtained. First, as was already known, many of the objectives were written in general rather than specific terms. Second, the objectives often contained vague behavior descriptions. Their clear meanings could not have been obtained without an examination of the instructional processes. Third, and perhaps most important, 34 percent of the objective/instructional process pairs were focusing on different aspects of the same general behavior.

Many of the objective/instructional process pairs focused on differing task formats or differing modalities. Several of these pairs had totally different behaviors or had differing reading content. A total of six (6) general categories were defined as follows:

Difficulty of Task - Skills presented in isolation vs. sentences, use of examples and non-examples.

Key Terms - Key instructional terms, vocabulary not presented or used in different ways in the objective and instruction.

Different Behaviors - Differing behaviors, no instruction using the behavior or more instruction beyond requirements of behavior.

Differing Mode of Behavior - Use of differing modalities in task such as verbal vs. visual.

Differing Content - Objective content not covered in the instruction.

Differing Format of Task - Presentation of tasks in a way that might interfere with the demonstration of the behavior.

Table 1 presents the results of the matching of the objectives and instructional processes. It can be seen that of the 470 objectives evaluated, 161 or 34% had instructional processes which did not match the objectives.

The table also identifies the non-matching objectives by type of non-match and basal level. The largest category of non-matching was that of Different Behaviors, accounting for 42% of the non-matching objectives. The second largest category was Difficulty of the Task, which contained 29% of the non-matching objectives. Differing Content contained 16% of the non-matching objectives. The remaining 13% were distributed over the other three categories.

Table 1

Frequencies of Objective/Instructional Processes
Non-Matches Types of Non-Matches, Matches and Total Objectives
by Basal Level and Total

Match Status	Level A	B	C*	D	E	F	G	TOTAL
TYPE OF NON-MATCH								
Difficulty of Task	29	8		3	9	0	0	46
Key Terms	4	1		0	4	2	3	14
Different Behaviors	8	3		8	9	12	27	67
Mode of Behaviors	0	2		0	2	0	1	5
Differing Content	4	1		3	8	7	2	25
Differing Format of Task	0	0		0	4	0	0	4
NON-MATCH TOTAL	45	15		29	33	21	33	161
MATCH	26	36		54	77	79	37	309
GRAND TOTAL	71	51		68	110	100	70	470

*Not Yet Completed

IMPLICATIONS

The results obtained in this study could, of course, be related to the basal series being used. Conversations with district and university personnel familiar with other basals indicate that the results would most probably be the same with any other basal series. In fact, independent review of several basal series is being conducted using the same methodology. The preliminary findings of these reviews have revealed similar problems.

If a reading CRT were developed using only the content domains of behaviors chosen from the basal series, then the validity of the item-based series, on about one-third of the behaviors, would be highly questionable. Clearly, if one is to obtain the best definition of behavioral domains, all aspects of the behavior including the instructional circumstances should be considered. While not always accessible, instructional information should be included in the domain definitional process.

APPENDIX A

TITLE I CRT INSTRUCTION/AGREEMENT WORKSHEET

Date _____

TITLE I CRT Instruction/Agreement Worksheet

Objective _____ Raters _____

Instruction	Testing
___ Content (NA)*	___ Content (NA)*
___ Directions	___ Directions
___Visual Demonstrations	___Visual Demonstrations
___Auditory	___Auditory
___Reading	___Reading
___ Task	___ Item
___Read	___Read
___Write	___Write
___Listen	___Listen
___ Taxonomy	___ Taxonomy
___Knowledge	___Knowledge
___Identification	___Identification
___Recognition or Recall	___Recognition or Recall
___Comprehension	___Comprehension
___Application	___Application
___Analysis	___Analysis
___Synthesis	___Synthesis
___Evaluation	___Evaluation
___ Response	___ Response
___Oral	___Oral
___Match	___Match
___Recall	___Recall
___Generative	___Generative
___Written	___Written
___Match	___Match
___Recall	___Recall
___Generative	___Generative
___Mark	___Mark
___Kinesthetic	___Kinesthetic
___Match	___Match
___Recall	___Recall
___Generative	___Generative

Comments _____

* Not Agree - Please make notes in the Comments Section explaining the reason for non-agreement.

4/19/82

APPENDIX B

CRT - OBJECTIVE/POSTTEST ITEM/INSTRUCTION AGREEMENT SHEET

CRT - Objective/Posttest/Item/Instruction
Agreement Sheet

Objective number:

_____

Agreement Non-Agreement	1 and 2	2 and 3	1 and 3
1. Objective:			
2. Posttest Item:			
3. Instructional Process:			

REFERENCES

Berk, R.A. A comparison of six content domain specification strategies for criterion-referenced tests. Educational Technology, 1980, 20, 49-52.

Ebel, R.L. Must all tests be valid? American Psychologist, 1961, 16, 640-647.

Glaser, R. Instructional technology and the measurement of learning outcomes. American Psychologist, 1963, 18, 519-521.

Guion, R.M. Content validity - The source of my discontent. Applied Psychological Measurement, 1978, 15, 321-327.

Hambleton, R.K., Swaminathan, H., Algina, J., Coulson, D.B. Criterion-referenced testing and measurement: A review of technical issues and developments. Review of Educational Research, 1978, 48, 1-47.

Linn, R.L. Issues of criterion-referenced measurement. Applied Psychological Measurement, 1980, 4, 547-562.

Mager, R.F. Measuring Instructional Intent: or Got a Match? Belmont, Ca.: Lear Siegler, Inc./Fearon Publishers, 1973.

Popham, W. James, Domain Specification Strategies. In Berk (Ed.), Criterion-Referenced Measurement: The State of the Art. Baltimore, Md: Johns Hopkins University Press, 1980.

Popham, W.J., Husek, T.R. Implications of criterion-referenced measurement. Journal of Educational Measurement, 1969, 6, 1-9.

Scandura, J.M. Structional approaches to behavioral objectives and criterion-referenced testing. Educational Technology, 1977, 17, 20-25.