DOCUMENT RESUME

ED 373 053 SP 035 399

AUTHOR Vitali, Gary J.
TITLE Factors Influencing Teachers' Practices in an Assessment Driven Reform.
PUB DATE Jul 94
NOTE 35p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Educational Assessment; *Educational Policy; *Educational Practices; Educational Principles; Elementary School Teachers; Elementary Secondary Education; *Evaluation Methods; Objective Tests; *Performance Factors; Secondary School Teachers; Teacher Attitudes; Teacher Behavior; Teacher Characteristics; *Teaching Methods
IDENTIFIERS *Kentucky Education Reform Act 1990; Performance Based Evaluation

ABSTRACT

This empirical study was conducted to explore both the influence of assessment on teachers' classroom practice during an assessment dri[illegible] educational reform, and the principle underlying assessment driven instruction, i.e., "what you test is what you get." Factors targeted at the level of assessment were performance-based or objective paper-and-pencil methods. Factors targeted at the level of the teacher were demographic variables, teachers' opinions of assessment, the assessment method, teachers knowledge of assessment, and teacher efficacy. In the Fall of 1992, self-report data were collected from 117 regular classroom teachers. According to the data analysis, the research did not support the "what you test is what you get" assumption. Rather, it was determined that teachers' classroom practices were influenced by methods of assessment, teachers generally taught toward the method of state-wide assessment most congruent with their existing classroom practices, and the feasibility of assessment driving teachers' classroom practices varies as a function of several factors endogenous to the teacher as well as to the assessment. Implications for educational policy-makers who plan assessment-driven reforms and statistical tables complete the document. Contains 47 references. (LL)

Factors Influencing Teachers' Practices in an Assessment Driven Reform

Gary J. Vitali

Send correspondence to:
Gary J. Vitali
3711 Delaney Ferry Rd.
Versailles, Kentucky 40383
Telephone: (606) 873-5650

July 1994

Factors Influencing Teachers' Practices in an Assessment Driven Reform

Gary J. Vitali

This was an empirical study that explored the principle underlying tenant of assessment driven instruction-- that what you test is what you get (Guskey, 1994; Vitali, 1993). Assessment's role in many present day educational reforms has changed from a mechanism to measure the attainment of desired ends to a means by which to achieve desired ends. Within an assessment driven educational reform teachers are expected to alter their existing curriculum and instruction practices to accommodate the assessment(s) that are imposed by policy makers.

To date the use of assessment to drive educational reform has not generally been efficacious (Brown, 1990; Sheppard, 1992). Yet the use of assessment to drive reform is escalating. It is becoming exceedingly important to better understand specific factors that influence assessment's effect upon teachers' classroom practices.

This research focused on the influence of assessment on teachers' classroom practice during an assessment driven educational reform. During the period of this study, the Kentucky Educational Reform Act of 1990 (KERA) was to employ both performance-based and objective paper-and-pencil assessments as the means by which to drive educational reform. Therefore, KERA presented a historically unique opportunity to explore the influence of nature of assessment upon teachers' classroom practices.

FACTORS EXPLORED

Assessment Means: Essential Differences

In an educational reform the nature of the innovation employed to facilitate change is a critical consideration. Within *KERA*, school performance with respect to student achievement is measured by two different types of assessment: objective paper-and-pencil and performance-based. Although these two types of assessment will be employed towards a common end, improved student achievement, they differ significantly at many levels.

Objective paper-and-pencil tests are often referred to as "indirect" assessments of student competencies (Cizek, 1991). This means of assessing student achievement has a long history in American schools. Examples of objective paper-and-pencil tests which have been used in Kentucky are the *Comprehensive Test of Basic Skills* (CTB/McGraw-Hill, 1982) and the *Kentucky Essential Skills Test* (Kentucky Department of Education, 1985). Objective paper-and-pencil assessments include a variety of response formats, e.g., multiple choice, true-false, and matching. However, the most commonly employed response format with an objective paper-and-pencil method of assessment is multiple-choice. Technologies that facilitate efficiency of scoring, reduce cost, and foster normative referencing of student performance are coveted by designers of this means of assessment. Critics note that the instrumental rationality upon which objective paper-and-pencil assessments are designed often serves the interests of method and efficiency at the expense of the interests of the students (Magnusson & Osborne, 1990).

Performance-based assessment differs from objective paper-and-pencil in many substantive ways. With performance-based assessment student responses to a test question are often more open, complex, and may have more than one correct answer per test question. Performance assessments tend to emphasize "judgements of process/action rather than product" (Garcia and Pearson, 1994). Oftentimes, performance-based assessments are embedded in the student's curriculum. A collection of the student's actual work may be gathered and then audited for desired characteristics. Student responses to assessment questions are often projects or portfolios (Frechling, 1991; Meyer, 1992). Performance-based assessments are generally designed to elicit student competencies in a given domain and not to reference a student's competency normatively (Wiggins, 1989). Popular literature often refers to this method of assessment as a "direct" assessment of the student's achievement. Critics of this means of assessment note the increased cost relative to time and money as well as difficulties with reliability.

The essential difference between objective paper-and-pencil and performance-based means of assessing student learning outcomes are pervasive and critical. Yet, there has been a relative dearth of study contrasting the influence of type of assessment upon teachers' practices. Given their marked differences, a logical deduction is that they would have a differential effect as mechanisms of change in an assessment driven educational reform.

Teachers: The Critical Consideration

In essence, policy makers intend to have teachers adopt the assessment innovations and align their instructional practices accordingly with the methods of assessment. Again, an essential underlying assumption of those who design test driven education reform is "what you test is what you get". A review of literature concerning the mechanics of teacher change reveals that change is a very complex process. Policies which simply mandate teacher change in an effort to accommodate a new policy or innovation are rarely successful (Fullan, 1991; McLaughlin, 1990).

Many policy experts have repeatedly determined that if an educational policy is to be successful, the current practice and beliefs of the teacher are necessary but not sufficient consideration for success (McLaughlin, 1990). Generally, teachers have negative attitudes and beliefs regarding external tests of their students' achievement (Marso and Pigge, 1989; Stiggins, 1991). However, teachers do appear to be receptive to alternative means of assessing student achievement. Recent studies suggest that most teachers have a preference for performance-based methods of assessing student achievement over traditional objective paper-and-pencil means (Stiggins, 1991; Wiggins, 1989). Over the past few years, the frequency of performance-based assessment methods in assessment driven reform efforts has increased rapidly.

Factors Which Influence A Teacher's Practice Towards Policy Innovations

In present day educational reforms the role of assessment has changed. Assessment methods employed as means of change have been referred to as assessment innovations (Cizek, 1991). A review of the literature revealed factors which are both exogenous and endogenous to teachers that bias the extent to which teachers adopt innovations.

Teachers' knowledge of the innovation can influence their perceptions of its' value and consistency with current practices. An inadequate knowledge of assessment may compromise a teacher's opinion of the practicality and costs involved with adopting the new assessments. Teachers' opinions regarding practicality and cost have been documented as factors which significantly influence adoption or change in practice (Doyle & Ponder, 1978-79; Hall & Loucks, 1978). Many current studies regarding teachers' knowledge or competencies with assessment reveal that in the aggregate they do not have the skills necessary to meaningfully understand or

modify current practices towards externally mandated assessments (Impara, 1991; Stiggins, 1991; Tittle, 1989).

Attributional characteristics such as teacher efficacy has been linked to the receptiveness in which a teacher implements innovations (Guskey, 1988). Teacher efficacy was the most powerful teacher factor related to instructional effectiveness in the Rand Corporations' Change agent study (McLaughlin & Marsh, 1978). However, to date there has been little attention to the relationship between teacher attributional characteristics related to student learning outcomes and the method of assessment employed to measure student learning outcome. Demographic factors become a logical consideration when studying factors exogenous to a teacher which may influence the extent to which externally imposed methods of measuring student learning outcomes effect change in teachers' classroom assessment and instructional practices.

Assessment driven education reform appears to be the dominant means of planned change in the 1990's. In an effort to specify more effective models of education reform, policy makers will need more insight into the relationship between teachers' assessment and instructional practices relative to the method of assessment employed by policy makers to assess student learning outcomes. This study will explore factors which influence the efficacy of method of assessment to affect teachers' classroom assessment and instructional practices. This study will specifically explore the relationship between a teachers' knowledge of assessment, sense of efficacy, opinion of assessment type, and demographic factors upon the alignment of classroom practice with state-wide objective paper-and-pencil and performance-based as employed in the early stages of *KERA*.

METHODS

This study employed a multivariate design. There are four predictor variables and one multivariate criterion variable. During the Fall of 1992, 117 teachers with regular classroom assignments from Central and Eastern Kentucky were randomly selected from a pool of 196 teachers who volunteered to participate in this study. Self report information was gathered via five questionnaires. Acuracy of teachers' responses on the criterion variable (teachers' instructional and assessment practices) was studied. Ten randomly selected teachers, five from each district, were observed on three observation visits for approximately four hours per visit. After performing classroom observations, review of lesson plans, and structured interviews with each teacher as well as interviews with their peers and students, I subsequently completed a teachers' instructional and assessment practices questionnaire for each of the ten teachers. There was a strong positive correlation between observed and reported teacher practices ($r=.86$; $p<.05$). This served to establish that teachers' responses on the questionnaire which addresses the criterion variable were reasonably accurate and representative of their actual practices.

Instrumentation

Data were gathered via five questionnaires. Descriptions of questionnaires are as follows:

Teacher Efficacy: Teacher efficacy will be assessed by the *Responsibility for Student Achievement Scale (RSA)*, (Guskey, 1981). The *RSA* contains thirty items designed to assess teachers' beliefs in their influence on student achievement. Questionnaire items are presented in an alternative-weighted format. Two subscales of scores are derived, one for efficacy regarding classroom success and the other for classroom failure. These subscales report unique constructs.

Test-retest reliability for the *RSA* was reported as moderately high, .784 for total R scores, .718 for R+ and .784 for R-. Two internal consistency indices were reported because the *RSA* contains two types of items, "those which sample beliefs in self-responsibility for either positive or negative classroom events" (Guskey, 1981). Coefficient alpha for the R+ items was .791 and for R- items .881.

Teacher Knowledge of Assessment: *Teacher Knowledge of Assessment Questionnaire*, (Plake and Impara, 1990) Part I consists of 35 items designed to assess proficiencies with the seven Standards for Teacher

Competency in Educational Assessment of Students. Questionnaire items were developed by the American Federations of Teachers (AFT), National Council on Measurement in Education (NCME), and the National Education Association (NEA). Questionnaire items represent the following seven domains: Choosing Assessment Methods, Developing Assessment Methods, Administering, Scoring, and Interpreting Assessment Results, Using Assessment Results for Decision Making, Using Assessment in Grading, Communicating Assessment Results, and Recognizing Unethical Practices. This questionnaire was employed in a collaborative project with the American Association of Colleges for Teachers in Education, AFT, NCME, and the NEA during a national survey to evaluate teacher assessment literacy (Plake, 1993).

Teacher Demographics: Demographics identified will be: teacher's county of current practice, grade taught, years of experience, and amount of education or training directly pertaining to objective paper-and-pencil or performance-based methods of assessment.

Teacher Opinions of Method of Assessment: This questionnaire consists of forty items designed to measure teachers' opinions of two methods of assessment of student achievement: objective paper-and-pencil and performance-based. Questionnaire construction closely followed the methods described by Kifer, (1977). Objective paper-and-pencil and performance-based methods of assessment are each targeted in twenty mapping questions. Facets of the questions are nested in four teacher practices: teaching, instructional planning and classroom structuring, assessment of student learning, and student preparation for mandated standardized testing of students. Elements within each facet are referenced to Doyle and Ponder's (1977-80) notions of instrumentality, congruence, and cost.

Teachers respond to a four point Likert scale which ranges from -2 (significant disadvantage) to +2 (significant advantage). This questionnaire was piloted during Spring, 1992. Internal consistency was computed for items subsumed under each method of assessment and then for all items. Coefficient alpha for the objective paper-and-pencil subscale was .94 and for the performance-based method of assessment subscale .90

Test-retest reliability was also calculated for each method of assessment subscale and for the entire questionnaire. The sample size for the test-retest reliability study was small (n=13). However, the reliabilities reveal a moderately high consistency of response: objective paper-and-pencil method of assessment .82,

performance-based method of assessment subscale .86 and for the entire questionnaire .83.

Criterion Variable: Teacher Classroom Assessment and Instructional Practices Questionnaire. This questionnaire consists of thirty items which comprise two, fifteen item subscales. Questionnaire construction closely followed the methods described by Kifer (1977). Facets for each subscale are assessment and instructional practices which are believed to facilitate a student's perfc mance on objective paper-and-pencil or performance-based methods of assessing student learning outcomes. Elements within each facet are referenced to teaching, instructional planning, classroom assessment, and standardized test preparation activities. Examinees' responses are percentage of time in which they perform each facet subsumed in each element.

Internal consistency was computed for each subscale and then for the entire questionnaire. Coefficient Alpha for the objective paper-and-pencil subscale was .91, performance-based subscale was .94 and the total was .41. The low coefficient alpha for the total questionnaire is believed to be an artifact of the strong negative correlations between the two subscales (r= -.76). This lends credibility to the notion that teaching towards an objective paper-and-pencil assessment may involve different practices than teaching towards a performance-based method of measuring student learning outcomes. Coefficient Alpha was recalculated during the fall, 1992 on 117 teachers. Coefficient Alpha for the objective paper-and-pencil subscale was .90, .92 for the performance-based subscale, and .57 for the total questionnaire. Correlation between subscales was (r= -.66). Test-retest reliability was calculated for each subscale and the questionnaire as a whole. The sample size for the test-retest reliability study was small (n=13). However, the reliabilities reveal a moderate consistency of response: teacher practices towards an objective paper-and-pencil assessment subscale was .86, teacher practices towards a performance-based method of assessment subscale was .77 and for the entire questionnaire .81.

ANALYSIS OF DATA

All grades within each school were represented. For the purpose of this study teachers were clustered into the following four groups: group 1; kindergarten-third grade; group 2: fourth-fifth grades; group 3: sixth-eighth grades; and group 4: ninth-twelfth grades. Table 4.1 depicts grade groups by district totals. The samples from both districts were fairly evenly dispersed across grade levels.

Table 4.1
Grade Group Distributions by School District

Group	Grade Levels	District A	District B
1	K-3	n=12	n=19
2	4-5	n=10	n=13
3	6-8	n=09	n=19
4	9-12	n=14	n=21
	Total	n=45	n=72

Table 4.2 presents a summary of descriptive statistics for the demographic variables: experience, college course hours pertaining to objective paper-and-pencil or performance-based assessment, and hours of staff development pertaining to standardized or performance-based assessment by district. Teachers' experience was not significantly different between districts. The number of hours that teachers received in college pertaining to objective paper-and-pencil or performance-based assessment was not significantly different between districts. Given that the term and the formal practice of "performance-based assessment" is a comparatively recent phenomenon in each county, it was not surprising that teachers in both districts reported very little formal exposure to this concept in college.

Teachers received the majority of their professional education pertaining to performance-based assessment via staff development in both districts. The variation in hours of staff development reported by teachers was influenced by factors operating at the district level and by the observation that the majority of teachers reported that staff development was generally self-selected within each district.

Table 4.2

Descriptive Statistics for Demographic Variables by District

	DISTRICT A			DISTRICT B					
	$\bar{X}$	(SD)	SEM	$\bar{X}$	(SD)	SEM	DF	t	p
Experience Teaching	11.42	(6.14)	0.57	13.70	(7.90)	0.73	(115)	1.67	.09
College course hours on:									
standardized assessment	4.80	(2.86)	0.26	3.15	(3.12)	0.29	(115)	1.20	.50
perf-based assessment	0.87	(1.63)	0.15	0.91	(1.91)	0.18	(115)	1.38	.20
Staff devel pertaining to:									
standardized assessment	5.68	(4.58)	0.54	3.92	(2.90)	0.29	(115)	1.32	.30
perf-based assessment	9.13	(7.93)	0.75	7.43	(4.04)	0.37	(115)	1.14	.61

Table 4.3 presents a summary of descriptive statistics for all variables by grade group. There were no statistically significant differences for demographic variables by grade group ($p<.05$). Notably, teachers in all grade groups report the vast majority of their college course work pertaining to assessment to be focused on objective paper-and-pencil assessment rather than performance-based assessment. Conversely, teachers in all grade groups report the majority of their staff development pertaining to assessment to be focused on performance-based assessment. This makes sense given the recent state wide emphasis on performance-based assessment. Given that teachers report no appreciable amount of formal education pertaining to performance-based assessment, the role of staff development becomes even more critical. Yet, on the average, teachers across all grades report approximately eight hours of participation in staff development pertaining to performance-based assessment.

There were significant differences in grades with teacher opinion of objective paper-and-pencil and performance-based assessment. Table 4.3a reveals that grades 4-12 do not significantly vary with respect to teacher opinion of objective paper-and-pencil assessment. Although, grades K-3 do differ from the rest of the

grades regarding teacher opinion of objective paper-and-pencil assessment. On the average teachers in grades K-1 reported a slightly negative opinion while teachers in grades 4-12 reported a moderately positive opinion of objective paper-and-pencil assessment.

Table 4.3b reveals that the means for teacher opinion of performance-based assessment also cluster into two grade groups K-3 and 4-12. However, teachers in grades K-3 report favorable opinions of performance-based assessment while the opinions of teachers in grades 4-12 are almost neutral. Given that there is not a significant difference is education or training by grade and the fact that performance-based assessment is a recent policy innovation for the teachers sampled in this study, the significant difference in teachers' opinions by grade may be explained by factors operative within curriculum and instruction. Perhaps the curriculum and instruction of teachers in grades K-3 tend to be more activity oriented while teachers in grades 4-12 tend to be content oriented. The performance-based assessment format may be perceived as more aligned with the classroom practices of teachers in grades K-3 and divergent from existing practices of teachers in grades 4-12.

There were no significant differences between the means of all four grade groups with respect to teacher knowledge of assessment. This is consistent with earlier findings regarding nonsignificant differences between grade groups with teachers' college education or staff development hours pertaining to objective paper-and-pencil and performance-based assessment. The amount of teacher education pertaining to either method of assessment is relatively small compared to the education received pertaining to curriculum and instruction. Given that both methods of assessment are designed to drive teacher practice across all grades in KERA there appears to be a critical need for increased teacher knowledge of assessment.

Positive teacher efficacy was significantly different ($p<.05$) across grades. Table 4.3c reveals that teachers in the lower grade groups have a higher sense of efficacy than teacher in the higher grade groups. The difference in teachers' sense of responsibility for student achievement by grade may be explained by the fact that teachers in secondary grades generally spend less time with individual students than do primary grade teachers. Also, the grade level difference may be explained in part by the fact that secondary teachers have older students with more firmly established patterns of learning which may not respond to change as readily as primary grade students (Guskey, 1982).

Teachers' practices towards objective paper-and-pencil assessment did vary as a function of grade. Table 4.3d reveals that once again the means grades K-3 are significantly different ($p<.05$) than grades 4-12. The differences between the average percentage of time that a teacher's classroom practices are aligned with objective paper-and-pencil assessments in grades 4-12 are nonsignificant ($p<.05$). Teachers in grades K-3 spend less time teaching toward an objective paper-and-pencil assessment than do teachers in grades 4-12.

Table 4.3e reveals that the grades cluster in a similar manner relative to the alignment of practice with assessment, but the direction of the alignment varies greatly for teachers in grade K-3. Teachers in grades K-3 spend more time teaching toward a performance-based method of assessment than do teachers in grades 4-12.

It makes sense that teachers in primary grades K-3 would not favor the objective paper-and pencil assessment format as student developmental skills necessary to respond to a closed response assessment format are not firmly in place during the early grades. The open student response format typically associated with performance-based assessment is more closely aligned with the criterion referenced assessment needs of primary grade teachers who are developing basic student learning skills as opposed to mastery of a specific domain of knowledge.

Table 4.3
Descriptive Statistics for Variables by Grade Group

	Grade Level							
	k-3 (n=31)		4-5 (n=23)		6-8 (n=28)		9-12 (n=35)	
	$\bar{X}$	(SD)	$\bar{X}$	(SD)	$\bar{X}$	(SD)	$\bar{X}$	(SD)
Experience Teaching	13.46	(7.35)	10.48	(6.27)	12.82	(6.91)	13.83	(8.25)
College course work in:								
obj. paper-&-pencil asmt.	3.70	(3.30)	2.87	(2.75)	3.86	(3.27)	3.54	(2.92)
perf.-based assessment	1.00	(1.75)	1.35	(2.31)	1.00	(2.14)	0.43	(0.98)
Staff development in:								
obj. paper-&-pencil asmt.	3.82	(3.90)	3.65	(4.67)	5.46	(5.85)	5.83	(6.45)
perf.-based assessment	7.90	(5.76)	7.87	(4.76)	8.36	(5.15)	9.06	(6.78)
Teacher Opinion of:								
obj. paper-&-pencil asmt.	-0.19	(0.81)	0.48	(0.82)	0.45	(0.78)	0.61	(0.85)†
perf.-based assessment	1.01	(.65)	0.17	(0.84)	0.33	(0.74)	0.13	(0.91)†
Tcher Knowledge of Asmt:	21.48	(5.20)	21.83	(4.91)	20.00	(4.08)	20.57	(4.39)
Teacher Efficacy:								
negative	57.77	(11.37)	55.59	(12.31)	52.89	(11.87)	57.53	(12.01)
positive	66.66	(9.90)	62.90	(12.05)	57.15	(9.83)	59.67	(11.09)†
Tcher Practice Towards:								
obj. paper & pencil asmt.	32.17	(12.55)	41.22	(16.92)	44.29	(18.00)	41.77	(18.21)†
perf-based assessment	60.53	(14.39)	41.45	(20.27)	39.38	(16.74)	34.27	(16.83)†

†Differs on post hoc (Duncan's) means test.
See post-hoc means comparison detail tables 4.3a-e.

Post-Hoc Means (Duncan's Multiple Range) Comparisons

Table 4.3a

Variable: Teacher Opinion of Objective Paper-and-Pencil Assessment

Duncan's Grouping	Mean	N	Grade Level
A	9.17	35	4
A	7.22	23	2
A	6.71	28	3
B	-1.77	31	1

Table 4.3b
Variable: Teacher Opinion of Performance-Based Assessment

Duncan's Grouping	Mean	N	Grade Level
A	15.19	31	1
B	4.96	28	3
B	2.52	23	2
B	1.94	35	4

Table 4.3c

Variable: Teacher Efficacy Positive

Duncan's Grouping	Mean	N	Grade Level
A	66.67	31	1
B A	62.90	23	2
B	59.67	35	4
B	39.38	28	3

Table 4.3d

Variable: Teachers' Practices Toward Objective Paper-and-Pencil Assessment

Duncan's Grouping	Mean	N	Grade Level
A	44.29	28	3
A	41.77	35	4
A	41.22	23	2
B	32.17	30	1

Table 4.3e

Variable: Teachers' Practices Toward Performance-Based Assessment

Duncan's Grouping	Mean	N	Grade Level
A	60.53	31	1
B	44.27	34	2
B	59.67	35	4
B	57.15	28	3

Table 4.4 presents a summary of descriptive statistics for teacher responses to the criterion variables: teachers' classroom practices toward objective paper-and-pencil assessments and performance-based assessments. Also included are the predictor variables: teachers' opinions regarding each assessment type, teachers' general knowledge of testing and assessment, and teacher efficacy.

There was a large range for both teachers' opinions of, and practices toward objective paper-and-pencil and performance-based assessment. For teachers' practices the unit of measurement is percentage of time spent teaching toward objective paper-and-pencil or performance-based assessments. Teachers who teach towards performance-based assessments on the average do so to a greater extent than teachers who teach toward objective paper-and-pencil assessments.

Teachers revealed a relatively low average score on the Teachers' Knowledge of Testing and Assessment questionnaire. There were 35 items and the average teacher answered approximately 21 correctly. This questionnaire was developed to identify teachers who require further professional development experience. Teachers demonstrated that they have only a modest knowledge of testing and assessment.

Table 4.4

Descriptive Statistics for Variables: Teacher Practices, Opinions, Knowledge of Assessment and Sense of Efficacy

(n=117)

	Mean	Median	(S.D.)	SEM	Range
Teacher Practices towards:					
Obj. Paper-&-Pencil Assessment	39.80	36.50	(17.10)	1.59	7.47-81.00
Performance-Based Assessment	46.80	47.33	(18.70)	1.73	4.00-82.67
Teacher Opinions of:					
Obj. Paper-&-Pencil Assessment	.53	.70	(1.29)	1.19	-1.80-1.50
Performance-Based Assessment	.73	.90	(1.30)	0.92	-1.60-1.80
Teacher Knowledge of Testing and Assessment Competency	20.91	21.01	(4.60)	0.43	10.00-31.00
Teacher Efficacy (+)	61.56	61.66	(11.2)	1.04	40.33-95.06
Teacher Efficacy (-)	56.10	55.33	(11.9)	1.10	29.27-83.27

The Criterion Variables

This study suggests that current classroom practices toward an objective paper-and-pencil assessment are different from classroom practices toward a performance-based assessment. Therefore, the analysis of the relationship between each predictor variable and the criterion will be structured such that teachers' classroom practices toward objective paper-and-pencil and performance-based assessments will be treated as separate variables.

Intercorrelations Between Predictor Variables

Table 4.5 presents a summary of correlations between all variables. There is a moderate inverse relationship between grade group and the percentage of time that teachers align their classroom practices with performance-based assessment ($r = -.32$; $p < .05$). This may be explained by the fact that unlike teachers in primary grades, teachers in the higher grades often teach multiple classes in different classrooms and the scoring and administration aspects of objective paper-and-pencil assessments is more expedient.

The correlation between teacher efficacy positive and teacher practices towards objective paper-and-pencil ($r = -.43$; $p < .05$) and performance-based methods of assessment ($r = .41$; $p < .01$) are both significant. However, the direction of the relationship varies with method of assessment. As a teacher's positive sense of efficacy becomes greater in magnitude the likelihood that classroom practices are aligned with objective paper-and-pencil assessments diminishes. Conversely, as a teacher's positive sense of efficacy increases so does the likelihood that classroom practices will be aligned with performance-based assessment. Given a review of the literature this makes sense. A positive sense of teacher efficacy is correlated with the likelihood that teachers will adopt and teach toward innovations (Guskey, 1986; McLaughlin, 1990). The negative correlation between a positive sense of teacher efficacy and diminished likelihood that classroom practices will be aligned with objective paper-and-pencil assessments also makes sense. Teachers' history of poor success with improving student learning as assessed by state-wide objective paper-and-pencil assessments would foster negative beliefs regarding this assessment method. As a teacher's sense of responsibility for student learning increases it is logical to conclude that assessment methods alleged not to be sensitive to their classroom practices would be

less favorable. Perhaps, it is not that teachers with a higher sense of efficacy prefer performance-based assessments, but rather that they are reacting against externally imposed objective paper-and-pencil assessments. A strong correlation exists ($r=.52$; $p<.01$) between a teacher's opinion of performance-based evaluation and the percentage of classroom practice aligned with this assessment method. However, there is a negative correlation ($r= -.42$; $p<.05$) between a teacher's opinion of performance-based assessment and the percentage of time spent teaching towards objective paper-and-pencil assessments.

At the time of the study teachers reported a marginal amount of college class work or staff development pertaining to performance-based assessment. Furthermore, on average, teachers demonstrated only a modest competency on the questionnaire designed to measure knowledge of assessment. Critics of objective paper-and-pencil assessment note that teachers have a history of dissatisfaction with this method of assessing student outcomes (Cizek, 1991; Shepard, 1989; Wiggins, 1989). Perhaps teachers in this study who prefer performance-based assessments are reacting against objective paper-and-pencil assessments as opposed to being pro performance-based.

The inverse relationship between knowledge of assessment and the percentage of classroom time aligned with objective paper-and-pencil appears to be complex. Logically, the more teachers know about assessment the more adaptive their classroom practices can be towards alternative assessment methods. Although objective paper-and-pencil assessments may be quicker to score and administer, the results may lack the richness of information the teacher desires.

A moderate negative correlation exists between a teacher's knowledge of assessment and the percentage of time that classroom practices are aligned with objective paper-and-pencil assessments ($r= -.38$; $p<.05$). The correlation between a teacher's knowledge of assessment and the percentage of time that classroom practices are aligned with performance-based assessment is not significant ($r=.16$; $p>.05$).

Table 4.5

Correlation Table

CORR	TchPrac-P	TchPrac-O	TchEff(-)	TchEff(+)	AsmtKnow	TchOp-P	TchOp-O	Stfdev-P	StfDev-O	ColHrs-P	ColHrs-O	Exper	Grade	Cnty
TchPrac-P	1.00													
TchPrac-O	-.66*	1.00												
TchEff(-)	0.34*	-0.36*	1.00											
TchEffic(+)	0.41*	-0.43*	0.36*	1.00										
AsmtKnow	0.16	-0.38*	0.13	0.12	1.00									
TchOp-P	0.52*	-0.42*	0.22*	0.25*	0.03	1.00								
TchOp-O	-0.27*	0.23*	-0.24*	-0.12	0.04	-0.55	1.00							
Stfdev-P	0.21*	0.18	0.14	0.05	0.21*	0.14	0.06	1.00						
Stfdev-O	0.05	0.07	0.03	0.09	-0.06	0.05	0.12	0.45*	1.00					
Colhrs-P	0.04	0.07	-0.09	0.09	-0.05	0.15	0.12	0.09	0.18	1.00				
Colhrs-O	-0.01	-0.02	-0.05	-0.07	-0.04	0.16	-0.17	0.26*	0.16	0.32*	1.00			
Exper	0.25*	-0.08	0.16	0.13	-0.04	0.10	0.07	0.09	0.15	-0.04	-0.07	1.00		
Grade	-0.32*	0.21*	-0.03	-0.26*	-0.11	-0.33*	0.31*	0.25*	0.15	-0.14	0.13	0.08	1.00	
Cnty	0.18	-0.06	-0.06	0.15	-0.17	0.06	0.07	0.00	0.04	0.00	-0.23*	0.16	0.02	1.00

* $p < .05$

Table 4.6 presents a summary of multivariate and univariate regression analyses for all predictor variables on the criterion variables teacher practices toward objective paper-and-pencil and performance-based assessment. There were no significant interactions ($p<.05$) between variables in this study. The models for objective paper-and-pencil and performance-based assessments explained 49% and 46% of the variance respectively. Type III sum of squares were employed in the calculation of F statistics.

Table 4.6
Multivariate Regression Analysis of Teacher Practices Toward Objective-Based and Performance-Based Assessment

Source	df	Parameter Estimates	Multivariate-F	Univariate F's	
				Objective-Based	Performance-Based
Constant	1	------	------	------	------
District	1	4.01	1.3	0.71	2.28
Grade	3	-3.03	3.97**	0.15	3.68**
Experience	1	0.37	1.43	0.00	2.64
Coll Hours-Obj.	1	-0.04	0.14	0.14	0.02
Coll Hours-Perf.	1	-0.09	0.56	1.85	0.00
Stf Dev Hrs-Obj.	1	-0.10	2.39	3.16	0.03
Stf Dev Hrs-Perf.	1	0.31	1.79	2.42	0.91
Asmt Knowledge	1	0.44	6.07**	12.66**	2.23
Opin-Obj Asmt.	1	0.84	.39	0.05	0.50
Opin-P-Based Asmt.	1	0.55	8.92**	10.61**	13.87**
Efficacy (+)	1	0.86	4.51**	9.91**	3.02*
Efficacy (-)	1	0.25	2.46	2.67	2.80
				Univariate Mean Squares	
Model	14			711.45	1348.04
Residual	98			263.76	284.50
Total	112				

* $p \leq .05$
** $p \leq .01$

Exploring the Relationships

A Summary

The first question explored in this investigation was "Do demographic factors influence teacher's classroom practices?" Table 4.6 presents a summary of multivariate and univariate regression results. Of the seven demographic predictor variables only grade had a significant multivariate F ($F=3.97$; $p<.01$). Examination of the univariate F's for each criterion variable revealed that there was no significant relationship between grade and teacher practices towards objective paper-and-pencil tests. However, there was a significant inverse relationship between grade and teacher practices towards performance-based assessment ($F=3.68$; $p<.01$).

Post hoc evaluation of the mean percentage of time that teachers in each grade group practiced toward one method of assessment versus the other revealed that teachers in grades K-3 practice toward performance-based assessment to a significantly greater extent ($p<.05$) than teachers in grades 4-12. Teachers in grades 4-12 revealed nonsignificant differences in the amount of time that they aligned their classroom practices with performance-based assessment.

Grade had a mild positive correlation ($r=.21$; $p<.05$) with teachers' practices toward objective paper-and-pencil assessments and a moderate negative correlation ($r= -.32$; $p<.05$) with classroom practices toward performance based assessment. It was hypothesized that the difference in the alignment of teachers' classroom practices with method of assessment could be explained by factors which were operative in the classroom. In this study grade differentially influenced teachers' classroom practices while experience, college education or staff development pertaining to objective paper-and-pencil and performance-based assessment did not. The statistically nonsignificant relationships between teachers' classroom practices and college education or staff development were believed to be an artifact of a restricted range of exposure to these educational programs

The second question explored was "Does teachers' sense of efficacy influence their classroom paractices?" The multivariate F for teacher efficacy positive was significant ($F= 4.51$; $p<.01$) but the multivariate F for teacher efficacy negative was not. Method of assessment has a differential relationship with the variable teacher efficacy. A teacher's sense of responsibility for student success (teacher efficacy positive)

has a significant influence upon the alignment of their classroom practices with performance-based or objective paper-and-pencil assessment methods. Both univariate F's for objective paper-and-pencil and performance-based assessment methods of teacher efficacy positive were significant ($F=9.91$; $p<.01$) and ($F=3.02$; $p<.05$) respectively.

A post hoc comparison of the mean degree of teacher efficacy positive by grade revealed that this attribution tends to diminish as grade level increases. In this study, primary grade teachers revealed a significantly higher degree of efficacy than middle or high school teachers. Teacher efficacy had a moderate positive correlation with teachers' classroom practices toward objective paper-and-pencil assessments ($r=-.43$; $p<05$). The relationship between teacher efficacy and performance-based assessment was similar in magnitude, but different in direction ($r=.41$; $p<.05$).

In part, the differential relationship between teacher efficacy positive and the alignment of classroom practice toward assessment may be explained by organizational factors. Teachers in primary grades tend to teach fewer students for a longer period of time than teachers in the secondary grades. Therefore, the sense of responsibility for student learning outcomes may be proportionally higher. Also, as a teacher's sense of efficacy for student learning increases it is logical to conclude that teachers would teach towards the method of assessment which they perceived would be more sensitive to their classroom realities. Given the activity driven nature of curriculum and instruction in the primary grades, performance-based assessment may be perceived as more consistent with existing practice than objective paper-and-pencil assessments, due to the fact that the latter method of assessment is more content oriented than process or activity oriented.

The third question explored was "Do teachers' opinions about the method of assessment influence their classroom practices?" The multivariate regression of teachers' practices on their opinions of objective paper-and-pencil assessments was not significant. In contrast, the multivariate regression for teachers' practices on their opinions of performance-based assessment was significant ($F=8.92$; $p<.01$). Once again assessment method is differentially related to the predictor variables.

The univariate regression of teacher practices towards objective paper-and-pencil assessment on their opinion of performance-based assessment was significant ($F=10.61$; $p<.01$) and the univariate regression of

teacher practices towards performance-based assessments on teachers' opinions of performance-based assessment was highly significant ($F=13.87$; $p<.01$).

A post hoc evaluation of the teachers' opinions of objective paper-and-pencil and performance-based assessments by grade revealed that teachers in grades K-3 differ significantly from teachers in grades 4-12 (see Tables 4.3a & 4.3b). Teachers in grades K-3 tended to have slightly negative opinions of objective paper-and-pencil assessments and moderately favorable opinions about performance-based assessment. In contrast, teachers in grades 4-12 had a slightly positive opinion of objective paper-and-pencil assessment and a relatively neutral opinion of performance-based assessment.

There were significant correlations between teachers' opinions of method of assessment and the extent to which they aligned their classroom practice with a method of assessment (see Table 4.5). There was a moderate correlation between teachers' opinions of objective paper-and-pencil assessment with practice toward this method of assessment. There was a negative correlation between opinion of objective paper-and-pencil assessment and practice toward performance-based assessment. The magnitude of correlation was much stronger between opinion of performance-based assessment and teachers' classroom practice toward method of assessment. Teachers with a favorable opinion of performance-based assessment aligned their classroom practice with performance-based assessment and did not tend to teach towards objective paper-and-pencil assessments.

A teacher's opinion of method of assessment is significantly related to the alignment of their classroom practice with objective paper-and-pencil or performance-based assessment. Perhaps factors operative within grade are directly related to teachers' opinions of method of assessment. It may be that a teachers opinion of the utility or practicality of one method of assessment versus the other greatly influences their opinion. Once again, factors operative at the level of the teacher influence the nature and magnitude of effect of assessment upon practice in an assessment driven reform.

The final question explored was "Does a teacher's knowledge of testing and assessment influence their classroom practices? The multivariate F for the predictor variable teachers' knowledge of assessment was significant ($F=6.07$; $p<.01$). A teacher's knowledge of assessment does influence their classroom practices relative to performance-based or objective paper-and-pencil assessments. The univariate F for the regression of

objective paper-and-pencil assessment on teachers' knowledge of assessment was highly significant ($F=12.66$; $p<.01$). The univariate F for the regression of performance-based assessment on teachers' knowledge of assessment was not significant ($F=2.23$; $p>.05$). Teachers' knowledge of assessment is related to teachers' classroom practices relative to objective paper-and-pencil assessments but not to performance-based assessments. Once again, method of assessment is differentially related to the factors targeted in this study.

A post hoc evaluation of the mean score on the knowledge of assessment questionnaire as a function of grade revealed nonsignificant differences ($p>.05$). In this study, teachers in all grades revealed a relatively modest level of knowledge of assessment. This finding is consistent with the limited amount of college course work or staff development teachers reported pertaining to assessment.

A moderate negative correlation exists between a teacher's knowledge of assessment and the percentage of time that classroom practices are aligned with objective paper-and-pencil assessments ($r= -.38$; $p<.05$). The correlation between a teacher's knowledge of assessment and the percentage of time that classroom practices are aligned with performance-based assessment was not significant ($r=.16$; $p>.05$).

The nonsignificant regression coefficient between a teacher's knowledge of assessment and the alignment of their classroom practices to performance-based assessment could be an artifact of the restricted range of knowledge of performance-based assessment. Teachers report less than one hour of college course work and an average of about eight hours of staff development pertaining to performance-based assessment (see Table 4.2). Given this modest amount of education there was a significant correlation between staff development pertaining to performance-based assessment and both assessment knowledge and teachers' practices towards performance-based assessment (see Table 4.5).

Summary and Dissussion

This study was exploratory in nature and sought to examine the influence of factors operating both at the level of assessment and the teacher, relative to teachers' classroom practices. Factors targeted at the level of assessment were assessment method. performance-based or objective paper-and-pencil. Factors targeted at the level of the teacher were demographic variables, teachers' opinions of assessment, assessment method, teachers'

knowledge of assessment, and teacher efficacy. This study has revealed that teachers' classroom practices are differentially influenced by method of assessment. There was a strong negative correlation between teacher practices towards performance-based and objective paper-and-pencil assessments.

The extent that teachers' classroom practices are aligned with performance-based or objective paper-and-pencil methods is influenced by many factors external to the assessment method. Factors significantly influencing the alignment of practice with assessment method were teachers' grade, knowledge of assessment, opinion of performance-based assessment and teacher efficacy positive. The relationship between each of these factors and teachers' practices towards performance-based or objective paper-and-pencil assessment methods tended to be inverse. That is, factors positively correlated with one method of assessment tend to be negatively correlated with the other. Teachers in grades K-3 and 4-12 tend to behave homogeneously relative to the factors targeted in this study, but differently from each other.

The underlying assumption of assessment driven educational policy that "what you test is what you get" was not supported by this study. The "what you get" or teachers' classroom practices are influenced by the method of assessment as well as several factors which operate at the level of the teacher. Teachers generally taught toward the method of state-wide assessment that was most congruent with their existing classroom practices. In conclusion, the feasibility of assessment to drive teachers' classroom practices varies as a function of several factors endogenous to the teacher as well as to the assessment.

Policy Implications

Educational policy makers who plan assessment driven reforms should acknowledge factors that are operative at the level of teacher as well as the assessment when designing a policy. In an effort to avoid large state-wide policies that drive only small classroom changes, factors operating at the level of a teacher must be valued and incorporated into policy designed to change existing teacher practices. Also, simultaneously mandating assessments of student learning outcomes which are divergent on many levels may serve to send mixed messages regarding what is valued or what is a desired outcome. The existing practice constraints upon teachers may not allow the time necessary to teach towards both assessment methods. Ralph W. Tyler (1949)

noted the importance of education based upon student outcomes but cautioned that our philosophy of education and desired goals should dictate the educational experiences that schools provide. Although Ralph Tyler's advice is over forty years old, it remains valid and apropos today. The influence of method of assessment upon desired goals should be an essential consideration of policy makers. The feasibility and compatibility of performance-based and objective paper-and-pencil assessments to drive teachers' classroom practices across all grades, subjects, and students has yet to be established.

If policy makers intend to drive teacher's classroom practices with an assessment innovation the.. teachers should have at least a functional knowledge of that innovation. Given that the teachers in this study revealed a modest and probably subfunctional understanding of assessment, it is logical to conclude that increased staff development efforts would be beneficial.

Secondly, organizational barriers to adoption of assessment such as grade need to be addressed. The effort to teach towards performance-based assessment for secondary teachers may be disproportionately larger than for primary grade teachers due to organizational factors. The implementation of an assessment driven reform may need to be grade sensitive.

Lastly, this study supports the notion that method of assessment critically impacts factors operating at the personal level of teachers, such as their opinion of assessment method. Dissonance between teacher opinion and policy mandates may serve to sabotage success. Experts have noted that teacher success may drive their beliefs or opinions, and that teacher change is often a . ~ and complex process (Fullan, 1991; Guskey, 1986). Policy makers may employ sustained and relevant staff development to facilitate teachers' knowledge of assessment and to change teachers' existing opinions about assessment methods.

REFERENCES

American Federation of Teachers, National Council on Measurement on Education, and National Education Association. (1990). Standards for teacher competence in educational assessment of students. Washington, DC: American Federation of Teachers.

Brown, P.R. (1990). Accountability in education. Policy briefs, Number fourteen. Far West Lab. For Educational Research and Development, San Francisco, Calf.: Office of Educational Research and Improvement (ED), Washington, DC.

Center for Research on Evaluation, Standards, and Student Testing (1990). Monitoring the impact of testing and evaluation innovations project: State activity and interest concerning performance-based assessment. (Technical Report) Los Angeles CA: Author

Cizek, G. J. (1991). Innovation or enervation? Performance assessment in perspective. Phi Delta Kappan, 73 (3) 695-699.

CTB/McGraw-Hill Book Co. (1982). Comprehensive test of basic skills. Monterey, CA: Author.

Doyle, W., & Ponder G. A. (1978-79). The practicality ethic in teacher decision making. Interchange, 8(3), 1-12.

Frechtling, J. A. (1991, Winter). Performance assessment: Moonstruck or the real thing? Educational Measurement: Issues and Practice, pp. 23-25.

Fullan, M. G. (1991). The new meaning of educational change. 2nd ed. New York: Teachers College Press.

Garcia G.E. & Pearson D.P. (1994) Assessment and diversity. Review of Research in Education, 20, 337-391.

Guskey, T. R. (1981). Measurement of the responsibility teachers assume for academic successes and failures in the classroom. Journal of Teacher Education, 32 (3), 44-51.

Guskey, T. R. (1982). Differences in teachers' perceptions of personal control of positive versus negative student learning outcomes. Contemporary Educational Psychology, 7, 70-80.

Guskey, T. R. (1986). Staff development and the process of teacher change. Educational Researcher, 15, (5), 5-12.

Guskey, T. R. (1987). Context variables that affect measures of teacher efficacy. Journal of Educational Research, 81 (1), 41-47.

Guskey, T. R. (1988). Teacher efficacy, self concept, and attitudes towards the implementation of instructional innovation. Teaching and Teacher Education, 4, 63-69.

Guskey, T. R. (1991, Spring). The wisdom of Ralph W. Tyler. Outcomes, pp. 6-7.

Guskey, T.R. (1994). What you test (assess) may not be what you get. Educational Leadership, 51, (6), 51-54.

Hall, G. E. , & Loucks, S. F. (1978). Teacher concerns as a basis for facilitating and personalizing staff development. Teachers College Record, 80 (1), 36-53.

Haney, W., & Madaus, G. (1989). Searching for alternatives to standardized tests: Whys, whats and whithers. Phi Delta Kappan, 70 (9) 680-685.

Herman, J. L. & Golan, S. (1992). Effects of standardized testing on teachers and learning -- Another look, (CSE Tech. Rep. No. 334). Los Angeles, CA: University of California, Center for the Study of Evaluation.

Impara, J. C., & Divine, K. P. (1991). Does interpretive test score information help teachers? Educational Measurement: Issues and Practice, 16-18.

Impara, J. C., & Plake, B. S. (1991). Unpublished test developed under a grant from the Kellogg Foundation to NCME and based on the Standards for Teacher Competence in Educational Assessment of Students developed by AFT, NCME and NEA (1990).

Kentucky Department of Education (1985). Manual for the Kentucky essential skills test. Frankfort, Ky.: Author

Kifer, E. (1977). An approach to the construction of affective evaluation instruments. Journal of Youth and Adolescence, 6 (3), 205-214.

McLaughlin, M. W. (1990). The Rand change agent study revisited: Macro perspectives and micro realities. Educational Researcher, 19 (9), 11-16.

McLaughlin, M. W. (1991). Test-based accountability as a reform strategy. Phi Delta Kappan, 73 (3) 248-251.

McLaughlin, M. W. & Marsh, D.D. (1978). Staff Development and school change. Teachers College Record, 80, 70-94.

Madaus, G. (1988). The influence of testing on the curriculum. In L. Tanner (Ed.), The politics of reforming school administration. London: Farmer Press.

Madaus, G. (1990). The distortion of teaching and testing: High stakes testing and instruction. Peabody Journal of Education, 65 (3), 29-46.

Madaus, G. F. (1991). The effects of important tests on students. Implications for a national examination system. Phi Delta Kappan, 73 (3) 227-231.

Marso, R. N. & Pigge, F. L. (1989). The status of classroom teachers' test construction proficiencies: Assessment by teachers, principals, and supervisors validated by analyses of actual teacher-made tests. A paper presented at the National Council on Measurement in Education, San Francisco. (ERIC Document: ED 306 283)

Magnusson, K. & Osborne, J. (1990). The rise of competency-based education: A deconstructionist analysis. The Journal of Educational Thought. 24, 1, 5-13.

Meyers, C. A. (1992, May). What's the difference between authentic and performance assessment? Educational Leadership, pp. 39-40.

Miller, M. H., & et al. (1990). A guide to the Kentucky Education Reform Act of 1990. Frankfort, KY: Kentucky Legislative Research Commission.

Odden, A. (1990, June-July). Educational indicators in the United States: The need for analysis. Educational Researcher, pp. 24-29.

Plake, B. (1993). Teacher assessment literacy: teachers' competencies in the educational assessment of students. Mid-Western Educational Researcher, 6, (i) 21-24.

Shepard, L. A. (1989). Why we need better assessments. Educational Leadership, 46 (7), 4-9.

Shepard, L. A, (April, 1992). Will national tests improve student learning? CSE Technical Report 342. Boulder, Colorado: University of Colorado.

Smith, M. L., Edelsky, C., Draper, K., & Cherland, M. (1989). The role of testing in elementary schools (Monograph). Tempe, AZ: Arizona State University, Center for Research on Evaluation, Standards and Student Testing.

Stiggins, R. J., & Conklin, N. F. (1988). Teacher training in assessment. (Report) Portland, OR: Northwest Regional Educational Lab.

Stiggins, R. J. (1991). Assessment literacy. Phi Delta Kappan, 73 (4) 534-539.

Stiggins, R. J. (1991). Facing the challenges of a new era of educational assessment. Paper to be published in Applied Measurement in Education.

Tittle, C. K., (1989). Validity: Whose construction is it in the teaching and learning contest? Educational Measurement: Issues and Practice, 8 (1), 5-13.

Tyler, R. W. (1934). Constructing achievement tests. Bureau of Educational Research, Columbus, OH: Ohio State University Bureau of Educational Research.

Vitali, G.J. (1993). Factors influencing teachers' assessment and instructional practices in an assessment driven educational reform. doctoral dissertation. University of Kentucky.

Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. Phi Delta Kappan, 703-713.

Wiggins. G. (1991). A response to Cizek. Phi Delta Kappan, 73(3) 700-703.