DOCUMENT RESUME

ED 355 511 CS 213 658

AUTHOR Bainer, Deborah L.; Porter, Frances
TITLE Teacher Concerns with the Implementation of Holistic Scoring.
PUB DATE Oct 92
NOTE 34p.; Paper presented at the Annual Meeting of the Midwestern Educational Research Association (Chicago, IL, October 14-17, 1992).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Educational Research; *Elementary School Teachers; Grade 3; *Holistic Evaluation; Primary Education; *Student Evaluation; *Teacher Attitudes; *Writing Evaluation
IDENTIFIERS Alternative Assessment

ABSTRACT

A study examined the concerns of five third-grade teachers over the first year of implementation of a district mandate to evaluate narrative writing using holistic scoring. Each teacher collected writing samples from six representative students. Teachers reported their responses throughout the year on two self-report instruments, and they were interviewed at the end of the year. Teacher concerns focused on the process of holistic scoring, the rubric itself, and communication with students and parents. Frustrations expressed by the teachers, however, revealed misunderstanding of the district mandate, inadequate inservice training, and confusion about writing assessment and the characteristics of "good" narrative writing. This hindered teachers from using holistic assessment effectively. Results suggest essential components of training if holistic scoring is to be efficiently, reliably, and validly applied in elementary school settings. (One figure is included.) (Author/SR)

Teacher Concerns with the Implementation of Holistic Scoring

Co-Authors:

Deborah L. Bainer, Ph.D.
Assistant Professor
Educational Theory and Practice
The Ohio State University, Mansfield
1680 University Drive
Mansfield, OH 44906

Frances Porter
Elementary School Teacher
Educational Theory and Practice
The Ohio State University, Mansfield
1680 University Drive
Mansfield, OH 44906

Paper presented at the annual meeting of the MidWestern Educational Research Association, Chicago, October 14 - 17, 1992.

ABSTRACT

The concerns of five third grade teachers following a district mandate to evaluate narrative writing using holistic scoring were examined over the first year of implementation. Teacher concerns focused on the process of holistic scoring, the rubric itself, and communication with students and parents. Frustrations expressed by the teachers, however, revealed misunderstanding of the district mandate, inadequate inservice training, and confusion about writing assessment and the characteristics of "good" narrative writing. This hindered teachers from using holistic assessment effectively. The results suggest essential components of training sessions if holistic scoring is to be efficiently, reliably, and validly applied in elementary school settings.

Teacher Concerns with the Implementation of Holistic Scoring

The task of teaching students to write provides a challenge and frustration for many teachers, as well as a vital experience for students. Writing not only enables students to communicate with their surroundings, but also provides them with an opportunity to strengthen their language and thought processes. Consequently, teachers search not only for strategies to better instruct their students in the craft of writing, but also for more effective methods of assessing students' writing skills.

Generally assessing students' writing serves two specific, informational purposes (Anderson & Lapp, 1988). First, assessment is informative, thus providing instructional feedback to the students about their work and enabling them to improve their skills. Second, assessment is evaluative, enabling the teacher to gain valuable insights into the students' understanding of concepts and mastery of writing skills. In this way, the effectiveness of the instructional program is also evaluated.

Research confirms that assessing writing provides a particularly formidable task for many teachers. Assessment problems are compounded when teachers must evaluate different types of writing: narrative writing that is highly imaginative, creative, and descriptive, and expository writing that is explanatory, clearly structured, and thorough. Traditional analytical methods of assessing student writing have been characterized as arbitrary, inconsistent, complicated, and dependent on the evaluator's definition of good writing. Allegedly, it fails to provide students with information about the specific skills

needed in order to write well (Haupt, 1990; White, 1984). Further, as Nickerson (1989) pointed out, traditional evaluation does not assess higher order cognitive functioning, such as creativity, which teachers are especially reluctant to criticize (Turbill, 1984).

Educators call for immediate research to develop and evaluate new assessment techniques. Assessment tools and systems that are efficient, reliable, and valid are yet to be adapted and widely implemented at the elementary school level. Efficiency requires that these methods be simple for teachers to use, flexible, and amenable to large scale usage (Huot, 1990; Turbill, 1984; Valencia & Pearson, 1987). In addition, these assessment techniques should establish objective standards to be followed to ensure reliability and to prevent "whimsical traditional grading" (Clark, 1987). Ideally, the assessment should focus on recognizing patterns in the students' writing (Johnston, 1987). Each standard used to assess the composition should be a description of the strengths and weaknesses of the work evaluated, like a diagnostic tool, allowing the students, parents, and other teachers to be aware of the quality of the paper (Indrisano, 1990). Finally, the assessment should reflect our understanding of how children learn (Valencia & Pearson, 1987) and eliminate extraneous variables that threaten the validity of an assigned score.

One of the biggest breakthroughs in writing assessment was the development of holistic scoring two decades ago (Huot, 1990; White, 1984). Holistic scoring, an efficient assessment approach, involves reading a writing sample quickly and making an overall judgment about its quality. Because holistic scoring assumes that all factors of writing are intertwined and equally important, the piece is evaluated as a coherent whole without analyzing its organization, mechanics, or

ideas (Cooper, 1983; Vacc, 1989). Yet holistic scoring is not based on unfocused judgment. To ensure reliable scoring, holistic evaluation asks students to respond to a carefully developed, precise prompt. Further, variability in the holistic scoring process is reduced by utilizing a specific numeric standardized scoring guide or rubric which provides direct statements or descriptors of papers at different points on the scoring scale. Anchor papers provide a specific example of each level of scoring along the rubric scale. Finally, reliability during holistic assessment is ensured with multiple independent scoring, usually by two readers.

Although this measurement system is still in its infancy, especially in elementary school settings, holistic assessment holds promise for teachers. Increasingly school districts, drawn by the efficiency of this method, are adopting holistic assessment. When holistic scoring has been applied in some situations, however, its validity has been challenged (Charney, 1984). Personal biases, content requirements, student syntactic and semantic usage, and extraneous features such as the neatness and length of the piece seem to effect a teacher's judgment (Turbill, 1984; Vacc, 1989). Research is needed to guide the implementation of holistic scoring methods in elementary school settings, for it is not known whether seasoned elementary teachers can adjust their attitudes and strategies to effectively and consistently evaluate writing in their classrooms using this approach or whether these threats to the reliability and validity of holistic scoring will be evident in elementary school settings as well.

The purpose of this study was to examine the impact of one school district's implementation of holistic scoring to assess writing. While the investigation explored the involvement of administrators, teachers, and

students in implementing the new assessment system, this report is delimited to a discussion of the reactions of the third grade teachers who were implementing rubric scoring for the first time. Specifically, the study sought to explore the following questions:

1. What concerns do teachers have related to the implementation of holistic scoring to evaluate writing?

2. How satisfied are teachers with using a rubric to holistically assess writing?

The Rubric Project

This study focused on the adoption of a policy by one school district requiring teachers to use a rubric to holistically evaluate their students' creative writing. The study involved third grade teachers in one lower elementary school (K-3) in an exempted village school district. While the school retained a small town character, it was proximal to an urban center and two universities with strong teacher preparation programs.

In line with district policy, the principal of the school implemented a four-year inservice program in writing instruction and assessment. During the first year, inservice education addressed how to teach and promote writing in the classroom. The inservice addressed the entire writing process. Teachers were taught to lead students to write drafts, peer conference, and revise before producing the final paper. Procedures for modeling creative writing and using writing portfolios were also presented to the teachers.

The inservice focus for the second year, during which this study was undertaken, was implementing holistic scoring to evaluate creative writing. The rubric adopted (Figure 1) was developed by a county-wide panel of teacher volunteers. The rubric was to be introduced to each

district as expediently as possible although no mandatory implementation date was set. However, the rubric was immediately adopted across the county to evaluate student writing proficiency, resulting in many schools attempting to train certified staff to use the rubric. With the urgency to use holistic scoring, the need to know how teachers implement holistic scoring and how successful the inservice was in preparing them to do so was apparent.

Figure 1

Method

Sample

The participants in this study were five third grade teachers, four females and one male. Each teacher had from 8 to 20 years of teaching experience. One of the teachers served on the district committee to develop the rubric, but none of the others had past experience with a rubric or holistic assessment.

The participating teachers selected six representative students from their classes whose writing assignments were evaluated during the study. That is, writing samples for the study were collected from a total of 30 students, one male and one female student of high, middle, and low ability from each of the five classrooms. Codes were assigned to participating teachers and student papers for confidentiality.

Instruments

Two instruments were designed to record teachers' reactions to holistic scoring. The Teacher Evaluation Form (TEF) asked teachers to evaluate samples of student writing and to self-report their reactions to

the process. The TEF provided a space to record the rubric score (1 through 5) given to each paper and asked teachers about their thinking and considerations while they decided the rubric score for each paper.

The Teacher Satisfaction Form (TSF) asked teachers to self-report their satisfaction with the rubric using a 5-point Likert Scale (very dissatisfied, okay, very satisfied). In addition, three questions were offered to probe participants' reactions to the evaluation tool and process, including how well the rubric addressed their assessment needs.

Procedure

All K-3 teachers in the building received training in the use of the rubric during two one-hour sessions after school in early November. During the first session teachers were acquainted with holistic scoring and the levels or standards of the rubric. Working in groups, teachers then selected anchor papers to represent the essential characteristics of each of the five levels of the rubric. They were cautioned that Level 1 and Level 5 papers, as outliers, are the hardest to find, but that it is often difficult to discriminate among Level 2, 3, and 4 papers to identify a representative Level 3 anchor paper. The importance of everyone interpreting the rubric levels in the same way and of using the anchor papers as models or guides when applying the rubric to writing samples were stressed.

After reviewing the characteristics of an anchor paper at each level of the rubric, the second inservice session engaged teachers in holistically scoring a sample of writing papers. Although encouraged to discuss the rubric levels, teachers were reminded not to talk about the specific papers being evaluated. Teachers were further cautioned to

read papers only once in order to provide holistic rather than analytical scoring. Each sample paper was read by two third grade teachers and the two separate rubric ratings were recorded on a tally sheet. If a discrepancy of two or more points was found between the teachers' ratings, a third teacher rated the paper. The individual ratings were averaged, providing a final rating.

At the end of the training, each third grade teacher developed a prompt for use in a writing assignment. For each of the following five months, one prompt was randomly assigned for use in a creative writing exercise in all five third grade classrooms. Each teacher was randomly assigned one month in which to submit a sample of six writing papers from the previously selected students in her class. These six papers were photocopied and distributed to the other third grade teachers for rubric scoring. Each month, teacher reactions were assessed using the TEF and TSF.

The following prompts were used for the creative writing papers in this study:

Prompt #1: You found a treasure chest, but you had to bury it. Write instructions for your family. Tell how to find your treasure.

Prompt #2: You have been chosen to go live on the moon. What will you take with you in your suitcase?

Prompt #3: Write about your "secret place." What do you do there? Where is it? It can be any place where you go alone.

Prompt #4: This class had to gather materials to solve some problem. They used the materials to construct their invention. In their writing assignment, they had to name the invention, tell what it was made of and what problem it solved for them.

Prompt #5: Walking in the woods and found a rabbit

After the fifth month of data collection, a final sample containing typed copies of six writing papers which previously had been assigned disparate ratings (i.e, at least 2 levels apart) was distributed to each third grade teacher for assessment. The student papers were typed to eliminate any bias in the evaluation caused by neatness or handwriting. Exit interviews with each teacher were also conducted and tape recorded with each teacher to probe the responses reported on the TSF and the TEF forms previously collected.

Analysis

Exit interviews and comments from the TEF and TSF forms were transcribed. Teacher responses were content analyzed by a jury of two to reveal primary patterns or categories of teacher concerns and levels of satisfaction related to the implementation of holistic scoring. Three broad categories of concerns emerged from the data. Teacher responses were then reread and classified according to these categories, then further notated based on areas of concern within the three categories.

Results

An analysis of the written and oral responses by the teachers regarding the implementation and use of a rubric to evaluate writing provided the following results with respect to teacher concerns and satisfaction with holistic assessment. In the following discussion, all teachers are referred to using feminine pronouns.

Teacher concerns

From content analysis, 187 statements of concern were identified. These statements recorded teacher concerns in three categories:

1. the process of holistic scoring (Scoring),
2. the rubric itself (Rubric), and
3. student and parent understanding of the rubric

(Students/Parents).

Teachers reported both positive and negative comments in each category.

1. Scoring. Teachers expressed concern with the process of holistic assessment in 94 (50.3%) of the concern statements. Teacher concerns with holistic scoring focused largely on the time required to evaluate papers using the rubric and the general usability and interpretation of the rubric.

Time. From the first month in which holistic scoring was implemented, four of the five participating teachers noted that it allowed for quick evaluation. This was attributed to the system's simplicity and to the fact that the evaluation gave an "overall view" of the students' work. As one teacher stated:

> You do not have to look for every
> mistake the student has made,
> therefore it takes you less time to go
> over the papers.

One teacher commented on the potential effect of this timeliness on the curriculum. She noted that teachers are more apt to engage their students' in writing if it doesn't take a lot of time to grade the papers.

Another teacher, however, found that holistic scoring slowed grading because "I had to concentrate on (the rubric) more than the way I have graded in the past." This teacher's need to refer to the rating scale frequently suggested that the rubric was never fully internalized by the teacher. After five months of implementation, this teacher conceded that grading "can be done quickly," but seemed unconvinced of the time efficiency of holistic scoring expressed by the other teachers.

Usability. Teachers agreed that using holistic scoring to evaluate writing papers was easy because the rubric provided specific points to follow, thus providing a "base on which to start." These baseline standards may have raised the expectations of the teachers regarding the quality of student papers. As one teacher noted,

> Before the rubric, everything that was even half-way decent was a satisfactory paper. It had to have lots of mistakes to get an "N".

Further, teachers noted that applying the rubric reduced the subjectivity inherent with evaluation. With the rubric, teachers were led to "look for the same things on everyone's paper" and to be more consistent in grading from one day to the next. The consistent use of the rubric by all third grade teachers was also seen as an advantage, because "everyone was measuring using the same standards."

Applying holistic scoring did not, however, entirely remove subjectivity from the teachers' evaluation. One teacher shared that she was more critical of a neatly written paper because the errors were easier to see, although she would not necessarily assign that paper a lower grade. Another teacher confessed being "more partial" to a paper that was neatly written because:

> I can read it; I don't have to think as hard about it. It is , clear right at the beginning what they are trying to say. Because usually the people that write the neatest are going to be the ones that will be the concisest and clearest in the assignment they have.

Further, some teachers acknowledged their subjectivity when they rated sample papers written by the children from their own class, whose handwriting they recognized. Two teachers seemed more favorable toward their own students' writing because they knew the students' abilities and could visualize what they meant, even if the students couldn't communicate it clearly in writing. Another teacher, however, reportedly graded her own students harder because, "I knew what they were capable of."

Interpretation. Interpreting the rubric while applying it to writing papers caused concern for the teachers. Most teachers expressed difficulty understanding and discriminating between some levels of the rubric. Some reported that they would "sit and agonize whether it was a 2 or 3," concluding that in the end it was "strictly a judgment call" and that "the more you read (the papers), the more you change your mind." Rather than wrestle with an individual interpretation of what the rubric levels meant when applied to difficult papers, one teacher resorted to assigning "2 1/2" to marginal papers, thus avoiding the struggle.

Other aspects of students' writing led the teachers to interpret rather than apply the rubric. One teacher shared her view that the "rubric means different things to different people." Creative papers seemed to pose special problems in grading.

> If the student does a creative job but doesn't really stay on the topic, then what? The teacher must decide which elements (of the rubric) are most important.

Another teacher secretly disagreed that the anchor papers

selected to represent Levels 4 and 5 deserved those ratings, "because they don't have what I would have looked for" in a strong writing paper. This teacher seemed to have developed her own rubric, at least subconsciously, and applied it instead of the adopted rubric to the writing samples. Although her interpretation of the rubric satisfied her own definition of good writing, the result was scores different from other teachers' ratings because of her subjectivity in interpreting the rubric.

2. Rubric. Teacher concerns in the Rubric category included issues that were inherent to the rubric's structure and, therefore, were outside of the control of the teachers. Rubric concerns were noted in 70 (37.4%) of the concern statements. Teacher concerns with the rubric focused on the specificity of the rubric and issues related to creativity.

Specificity. Teachers praised the rubric because it was concise and provided specific guidelines for evaluating student papers. In short, it was "a good measuring device," yet teachers seemed troubled with the "vague" wording at each level of the rubric. As one teacher reflected:

> Sometimes the words 'often, generally, usually,' etc. are hard to understand. For instance, 'generally uses correct spelling'. Does that mean misspelling one, two, or three words puts you in the lower category?

One teacher called for a "more realistic" rubric; another suggested that percentages, which teachers are more familiar with, should be included in the rubric.

> Most of our competencies have

> percentages on them. Like 80% of the time they use capital letters. A percentage would be more helpful than what it says here, because I never know, I mean, if you have five sentences, if you have three capitals is that good enough or four? Percentage would at least help because then you could go by how many sentences there are. It would be more specific but would not tie you down to certain numbers or things.

Further, teachers were concerned that certain mechanical aspects of good writing were not included in the rubric. Teachers observed that the rubric did not specifically address grammar and usage, the need for a theme or main idea, or the need for a conclusion and distinct beginning, middle, and end in the narrative. One teacher, however, dismissed the concern for considering the mechanical aspects of writing when evaluating with a rubric. If you want to grade for mechanics, she explained, "use another method."

Creativity. Beyond mechanics, throughout the five months of observation teachers wrestled with how to deal with creativity in writing when using the rubric. Noting that Levels 1, 2, and 3 in the rubric do not include creativity, most teachers were puzzled about how to evaluate papers for children who are very creative but "miss the boat on writing itself." Some teachers saw creativity as the "tie breaker" for marginal papers, creativity netting a higher rating for the student's paper. One teacher suggested that the reason for disparate ratings

being assigned to the same paper was the raters' varied reactions to creativity.

3. Students/Parents. The third area of concern for teachers centered around the students and parents. Concerns about students and parents were expressed in 23 (12.3%) of the concern statements reported by teachers. Teacher concerns were reported regarding how to provide feedback and meaningfulness for parents and students when utilizing holistic scoring.

Feedback. The teachers in the study agreed that students must be given specific feedback on their writing in order to improve. Further, they felt that third grade children are unable to proofread and find the mistakes in their papers unless the teacher has indicated those errors. From that perspective, using holistic scoring was "frustrating" for the teachers because it did not call for making corrections, did not allow teachers to edit the students' papers, and did not help students find their mistakes. As a result, one of the teachers who shared rubric scores with the students concluded that utilizing holistic scoring to assess student papers was "not meeting their needs." That is, as used by these teachers holistic scoring did not meet the informative purpose of evaluation.

Meaningfulness. The teachers in the study also agreed that the rubric numbers held little meaning for students and their parents, and this concerned them. The teachers shared few, if any, rubric scores with their students because "numbers do not tell them anything." Another teacher shared: "I explained (the numbers) to them, but I don't think they grasped it." One teacher wrestled with the issue of how to educate the parents and students as to the meaning of the numbers, fearing that the 1, 2, 3, 4, 5 rating system would be wrongly interpreted as A, B, C, D, F.

Teacher Satisfaction

The teachers self-reported being "satisfied" or "more than satisfied" with the holistic approach, but the satisfaction of all but one of the teachers wavered or diminished somewhat over the five months of the study. Discontent and frustration were evident in many of the teachers' comments, but this seemed to be related to issues surrounding the adoption of holistic scoring rather than to the system itself. Teacher discontent seemed to be related to four issues:

1. understanding the mandate to use holistic scoring,
2. understanding the inservice training,
3. understanding the assessment of writing, and
4. understanding "good" writing.

1. Understanding the mandate. Teachers' comments revealed that there was confusion regarding exactly what the district mandate required. Teachers were unclear about the scope of the adoption; that is, if holistic scoring was to be used to assess all writing or only creative writing. Further, they were confused about if and how they should use the rubric in their classrooms. Teachers were uncertain if they were to use the rubric to teach the writing process, or if holistic assessment should be limited to evaluating final papers for the students' portfolios or permanent files. Relatedly, teachers were unclear about who should see the ratings; whether or not the rubric system and ratings subsequently assigned to student papers were to be shared with the students and their parents. Throughout the study, teachers reiterated their desire to have someone tell them specifically what to do, or to tell them to sit down together and discuss how and when to use holistic scoring in their classrooms.

Because of the confusion about the directions of the mandate,

holistic scoring was implemented very differently across classrooms. Three types of teacher responses were obvious in the way they utilized the holistic assessment. The ambiguity resulted in avoidance by two teachers, who used holistic scoring only to rate papers that were part of this study. While one teacher confessed that she was "real lax" about teaching writing in her classroom and generally did not evaluate writing anyway, the other teacher expressed willingness to use the rubric in her classroom if only she knew how to implement it "properly." Throughout the study, she was in a quandary about how to educate the parents and students about the rubric and what the numbers mean, and how to use it to teach writing to the children.

Another teacher responded to the confusion about the mandate with resignation. She reportedly used the rubric in her classroom "just as it says" in spite of the fact that she had difficulty applying the rubric. Further, she was frustrated because she found that the rubric was "not helpful in getting students to learn to write." Why, then, had she resigned to using the rubric? Because it has been mandated. So, this teacher made it her goal to use holistic scoring in her classroom, feeling that "I have to do it because I have to become used to it." While another teacher wasn't certain that holistic scoring was to be applied to all creative writing assignments, she thought this requirement was "probably coming" but that, like other education fads, it would fade away after five or six years.

Finally, two teachers responded to the mandate by adopting holistic scoring in their classrooms, although that adoption was not without some confusion about the district's requirements. While one teacher used holistic assessment on only those papers destined for the students' permanent files, the other routinely applied the rubric to

writing papers diagnostically and for her own information. She then talked with those students who received low rubric scores (1 or 2) about what was wrong with their papers.

2. Understanding the training. Teachers' comments also suggested that the inservice training was ineffective in establishing the mindset necessary for successful implementation of holistic scoring. Throughout the study, teachers commented that using the rubric "would require us to retrain our minds" and to think differently when evaluating students' writing.

Specifically, three of the teachers involved in the study did not seem to understand or internalize the process of holistic scoring. Rather, they persisted in analyzing students' work in spite of the holistic approach required by the rubric. These teachers commented that it was difficult to look at papers holistically and to get away from looking for punctuation, capitalization, and other mechanical aspects of writing. Because of their analytical mindset, some teachers related "agonizing" experiences associated with assigning a rubric score. Others reportedly vacillated between adjacent rubric scores. Some teachers considered the holistic approach a matter of personal judgment, the rating scores thus showing differences in teacher expectations. Comments indicated that while these teachers endorsed having guidelines for assessment such as the rubric provided, they preferred their more concrete "grading systems with points off" to the holistic description of the writing provided by a rubric score. In short, their minds were not retrained to view writing holistically by the inservice experience.

Another concept which some teachers did not understand from the inservice was that the rubric should be applied, not interpreted.

Indeed, it is this aspect of a rubric that establishes the reliability of holistic scoring. One teacher, for example, disagreed with the standards in the adopted rubric and mentally created her own rubric which she applied to the students' papers. If one of the advantages of holistic scoring is that it saves time by eliminating the need to develop or interpret a grading system, this benefit was lost for this teacher and the reliability of the scoring process was jeopardized.

The observation that the teachers' minds were not entirely retrained to a new assessment approach by the inservice education was illustrated by one teacher's explanation of how she assigned rubric scores:

> ...the way I basically use the rubric and arrive at a number is probably basically how I also arrive at a letter grade...I just sort of read it and a letter grade would come to you like when you read it and one of these numbers comes to you.

3. Understanding the assessment of writing. Teacher discontent with holistic scoring also seemed related to their understanding of writing assessment. While some teachers in the study were confused about which writing papers to evaluate and how to evaluate them, others had resolved the conflict in various ways.

Specifically, two teachers expressed frustration that seemed to be rooted in their unresolved questions about the assessment of writing in general. One teacher found it "very difficult" to grade writing and avoided assigning either number or letter grades to writing unless "I have to put it on a report card; then I need a lot of grades." This teacher

disliked grading writing because it was "attitudinal" or subjective. For her, utilizing holistic scoring did not resolve this conflict because it didn't matter if numbers or letters were used: they were all grades. This teacher's approach to assessing writing tended to accommodate rather than resolve her conflicts regarding assessment. When applying the rubric, she avoided difficult evaluation decisions by assigning "1/2" ratings to marginal papers. Further, she advocated a "Satisfactory/Not Satisfactory" evaluation system in which any completed work would be awarded an "S" rather than a grade.

The second teacher's conflict regarding the assessment of students' writing centered on her expectations of pupils' writing. Confusion regarding how and what to assess arose from her belief that students cannot be both creative and mechanically correct. For this teacher, writing was "a lot of hard work for the teacher," whose role she defined as editing and pointing out mechanical errors for students because they are "unable to understand anything they are doing wrong." Because of the effort required and because of the confusion over what to look for, this teacher "didn't have them do much writing" and reported that "I just don't think I ever graded writing before." Her conflicts with writing assessment thus remained unresolved because she seldom engaged her students in writing and therefore avoided facing issues related to assessment.

In contrast, the remaining three teachers expressed little or no confusion about writing assessment. They had clearly determined how and when to evaluate creative writing, although their reasoning differed somewhat. One teacher frequently engaged her students in writing, but did not grade everything the students wrote. Further, she did not consider mechanical aspects of writing because "misspelled

things are fine because that is part of their creativity." Why did the teacher approach assessment in this way? Reportedly, "because that is how I was taught in college," over a decade ago. The two other teachers also clearly differentiated the purposes of writing and assessment in their own minds. They used holistic scoring only for creative writing assignments and papers destined for permanent student folders. These teachers used other assessment methods to evaluate expository writing papers or to focus on the mechanical aspects of writing, and neither teacher considered the rubric useful or appropriate for teaching students how to write.

4. Understanding "good" writing. Closely linked to some teachers' questions about how to assess writing was confusion over exactly what "good" narrative writing was. Specifically, it was apparent that there was no consensus among the teachers about the purpose of the writing experiences (to create, explain, or persuade), although most had determined goals in their own minds. The teachers' remarks separated them along a continuum between two "camps," articulated by one teacher as a concern for creativity only as opposed to a concern for both creativity and structure in the writing. While two teachers shared that narrative writing should emphasize creativity only, and that it was "unrealistic to expect both" from students, another teacher was "torn between the two" positions. Her emerging philosophy of writing suggested an emphasis on creativity wedded with attention to form:

> I would rather have kids write what
> they think and really get it all out and
> stop worrying so much about how they
> are saying it and everything else.
> Later, after writing has started, they

can be concerned about mechanics.

In contrast, other teachers asserted that they disagreed with the conflict between creativity and structure because structure is essential to creativity. One teacher explained her viewpoint about how knowledge of the mechanical aspects of writing provided structure and enabled students to write creatively.

> If you just tell kids to write, some of them would just be lost. They have no idea what to do, they are frustrated, they feel inferior with it so they don't do a good job. Sometimes you have to talk about structure a little. Structure gives them something to hang on to so they can get the job done.

The lack of clarity about the purpose of writing was especially evident when the teachers developed and talked about the writing prompts. One teacher discussed the relationship between the nature of the prompt and the level of creativity in students' papers. Citing students' responses to the prompts used in this study as an example, the teacher pointed out that her carefully structured prompt (Prompt #4: This class had to gather materials to solve some problem. They used the materials to construct their invention. In their writing assignment, they had to name the invention, tell what it was made of and what problem it solved for them.) generated papers "rich in individuality and creativity," whereas a less structured prompt (Prompt #5: Walking in the woods and found a rabbit) resulted in papers that were far less creative and "didn't seem to say anything." Interestingly, the author of the less structured prompt (Prompt #5) agreed that the prompt is

important in stimulating creativity. Because of her past experience with prompts that asked students to write about topics that were unfamiliar to them, she seemed to associate structure with difficult topics, and with expository rather than creative writing. Her position that "if you are going to get creativity, it's got to be (a prompt) that the kids can write about" was supported by other teachers. In her attempt to avoid writing a difficult, expository prompt, however, she produced a prompt that, according to other teachers, was too vague to stimulate creativity among the students in this study.

The variety of structure and content in the prompts written for this study may reflect the teachers' differing understandings of what narrative writing is and how to prompt students to write in the desired mode. Many of the prompts, in fact, seem to call for expository rather than narrative writing. And, if the teachers lack a clear focus on what they are looking for in writing papers or on what creativity is, how do they know what kind of prompts to write?

Conclusions and Discussion

This study identified the specific problems and frustrations experienced by teachers when holistic scoring was implemented in an elementary school setting. The problems teachers expressed about holistic scoring, about the structure of the rubric, and about explaining holistic scores to students and parents extend our understanding of the concerns associated with the implementation of holistic scoring at the elementary level. Interestingly, the concerns are similar to those expressed when holistic assessment was implemented in other situations (Charney, 1984; Turbill, 1984; & Vacc, 1989). Further, the results show that the product orientation cited by Huot (1990) as inherent with holistic scoring was a source of frustration and did limit the ability of

the elementary teachers to make informed decisions about how to use holistic assessment in their classrooms for writing assignments and instruction. These problems threaten the validity and reliability of holistic scoring in this elementary school setting, making the scores assigned by teachers in this study questionable.

From his experience with holistic scoring in higher education, White (1984) cautioned that holistic scoring, like many other "complex and handy" educational concepts, is frequently misunderstood, misused, and misinterpreted. This also seems to be the case with the implementation of holistic scoring by the elementary teachers observed in this study. In spite of the inservice efforts, it cannot be stated with confidence that the teachers truly understood or appropriately used holistic scoring, although they may have paid lip service to a rubric adopted by the district. The teachers' concerns and frustrations remind us that, in our eagerness to take advantage of the efficiency offered by holistic scoring, we often neglect other procedures that are essential to the effective implementation of holistic assessment in elementary school settings. This study, then, provides recommendations to guide inservice education if holistic scoring is to be effectively implemented, at least at the elementary level.

First, teachers must be taught to clearly recognize <u>when</u> it is appropriate to use holistic scoring. Teachers must realize exactly what holistic scoring can and cannot do. Holistic scoring, after all, merely rank-orders students along a rubric continuum. It cannot tell the teacher much about a student, nor can it tell students specifically how to improve their writing skills. It provides no insights into the context of the child's writing endeavor, the specific weaknesses which led to the assigned rubric score, or the child's other writing strengths not

assessed by the rubric. As White (1984) noted, holistic scoring gives the teacher a score for the student, but the teacher may want a profile. That is, at least when used with elementary students, holistic scoring is more evaluative than informative as an assessment system. This realization should lead teachers to look for patterns when applying holistic scoring, as suggested by Johnston (1987), rather than to agonize over scores for marginal papers as reported by teachers in this study. Elementary teachers, then, need to be taught how to make assessment decisions. They need to be able to recognize when to use holistic scoring and when other means of assessment would provide the desired information.

Second, teachers must clearly understand how to use holistic scoring in order to ensure valid scores. Because holistic scores are not absolute values, they are not generalizable (Huot, 1990; White, 1984). That is, a holistic score depends on the nature of the rubric applied, the difficulty and focus of the prompt used, and the particular group of students being assessed. Holistic scores only have meaning, then, in the light of these variables. This suggests that when elementary teachers involved with holistic scoring are being inserviced, they first must clearly identify the outcomes they want to assess. Next, they must develop a rubric that clearly and specifically addresses those outcomes at each level. If a higher order cognitive skill such as creativity is a desired outcome, then each level of the rubric must reflect varied degrees of that skill. This limits the usability of the rubric to the situation or task for which it was developed, and ensures the validity of scores thus assigned. Further, teachers must be taught to write concise prompts with sufficient structure and content focus to stimulate the desired pupil performance, whether it be expository or narrative

writing.

Third, elementary teachers must be provided with initial and ongoing training in order to ensure the reliability of holistic assessment. Vacc (1989) found that an initial training program was not enough to change many teachers' attitudes toward writing and writing evaluation, and that those ideas affected how teachers evaluated writing samples. The same was true in this study. This suggests that the training for holistic scoring needs to go beyond the mechanical aspects of assigning and tallying rubric scores to probing teachers' thinking when they assigned the rating. Until the teachers' minds have been retrained, holistic scoring cannot be carried out reliably. In addition, training must continue until teachers agree on the anchor papers and the characteristics of each score along the rubric; evidence that they have internalized the rubric. Continued communication among teachers and periodic monitoring throughout the scoring process further guards against scoring drift. As White (1984) noted, this "calibration" of teachers is essential to ensure reliability, and initial and ongoing training must allow sufficient time for this.

Finally, the training must convince participating elementary teachers of the value of holistic scoring. That is, there are more benefits to holistic scoring than efficiency and timeliness. From his experience with holistic scoring, White (1984) shared that the social interaction of colleagues working together, even in a tedious evaluation effort, led teachers to reconsider their instruction and evaluation practices and to value holistic scoring. He reported that a "sense of community" developed among those engaged in holistic scoring, and that this communal spirit was essential to enforcing the standards of the rubric and thus ensuring reliability in scoring. This suggests that

affective aspects of training, such as team building, cannot be neglected. Indeed, establishing communication networks among teachers during training is essential to the successful implementation of holistic scoring. While impersonal, mechanical training can create an uncomfortable environment in which elementary teachers may comply with mandated holistic assessment, sensitive leadership and collaborative training can lead to cooperation and purposeful scoring that can impact teaching and assessment practices, and ultimately student learning.

While this study identifies problems and issues related to the implementation of holistic scoring by elementary school teachers, further research is clearly needed to fully understand the interaction between training procedures and participating teachers. Factors that contribute to teachers' holistic evaluation of writing samples need to be identified and explored. Further, training procedures that effectively retrain teachers' minds and establish a supportive community among teacher-evaluators must be developed and validated if holistic scoring and other breakthroughs in assessment practices are to be effectively applied in elementary school settings.

References

Anderson, P. S. & Lapp, D. (1988). Language skills in elementary education. New York: Macmillan.

Charney, D. (1984). The validity of using holistic scoring to evaluate writing: A critical overview. Research in the Teaching of English, 18, 65-81.

Clark, R. P. (1987). Free to write: A journalist teaches young writers. Portsmouth, NH: Heinemann.

Cooper, C. R. (1983). Holistic evaluation of writing. In C. R. Cooper, L. Odell (Eds.). Evaluating writing: Describing, measuring, judging. Urbana, IL: National Council of Teachers of English.

Gray, B. G. (1990). Easing the paper load: How to respond to writing. Professional handbook for the language arts: World of language. Morristown, NJ: Silver Burdett and Ginn.

Haupt, P. (1990). Holistic evaluation. Professional handbook for the language arts: World of language. Morristown, NJ: Silver Burdett and Ginn.

Huot, B. (1990). Reliability, validity, and holistic scoring: What we know and what we need to know. College Composition and Communication, 41 (21), 201-213.

Indrisano, R. (1990). On informal assessment. Professional handbook for the language arts: World of language. Morristown, NJ: Silver Burdett and Ginn.

Johnston, P. (1987). Teachers as evaluation experts. The Reading Teacher, 40 (8), 744-748.

Nickerson, R. S. (1989). New directions in educational assessment. Educational Researcher, 18 (9), 3-7.

Turbill, J. (1984). No better way to teach writing! Portsmouth, NH: Primary English Teaching Association.

Vacc, N. N. (1989). Writing evaluation: Examining four teachers' holistic and analytic scoring. The Elementary School Journal, 90 (1), 87-95.

Valencia, S. and Pearson, P. D. (1987). Reading assessment: Time for a change. The Reading Teacher, 40 (8), 726-732.

White, E. M. (1984). Holisticism. College Composition and Communication, 35 (4), 400-409.

Figure Caption

Figure 1. Assessment Scale for Third Grade Writing

Assessment Scale

Grade 3

Level 1:
1. Lacks use of correct capitalization and punctuation.
2. Lacks use of correct spelling.
3. Uses incomplete sentences.
4. Lacks paragraph organization.
5. Lacks a complete thought.

Level 2:
1. Shows deficiencies in use of capitalization and punctuation.
2. Seldom uses correct spelling.
3. Seldom uses complete sentences.
4. Often lacks paragraph organization.
5. Lacks focus.

Level 3:
1. Often uses correct capitalization and punctuation.
2. Generally uses correct spelling.
3. Usually uses complete sentences.
4. Attempts to use paragraph organization.
5. Often focuses on topic.

Level 4:
1. Usually uses correct capitalization and punctuation.
2. Usually uses correct spelling.
3. Uses complete sentences with some variety.
4. Usually uses organized paragraphs.
5. Usually focuses on topic.
6. Shows some creativity.

Level 5:

1. Uses correct capitalization.
2. Uses correct spelling.
3. Uses good sentence structure and expanded vocabulary.
4. Uses organized paragraphs with detail sentences.
5. Focuses on topic.
6. Uses creativity.