

UC Davis

Journal of Writing Assessment

Title

A prominent feature analysis of seventh-grade writing

Permalink

<https://escholarship.org/uc/item/31n062h1>

Journal

Journal of Writing Assessment, 4(1)

Authors

Swain, Sherry S.
Graves, Richard L.
Morse, David T.

Publication Date

2011

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer reviewed

A prominent feature analysis of seventh-grade writing

by Sherry S. Swain, Richard L. Graves, and David T. Morse

Abstract

The present study identified the characteristics of seventh-grade writing produced in an on-demand state assessment situation. The subjects were 464 seventh graders in three middle schools in the southeastern United States. The research team included 12 English language arts teachers. Results of the analysis yielded some 32 prominent features, 22 positive and 10 negative. The features were correlated with state assessment scores, which ranged from 1 to 4. Of the 22 positive features, 14 correlated positively with the assessment scores. Of the ten negative features, 8 correlated negatively with the assessment scores. The study also found 108 statistically significant ($p < .001$) intercorrelations among the features. From the features themselves, a formula was devised to create a prominent feature score for each paper, the scores ranging from 3 to 21. The prominent feature scores were also significantly correlated with assessment scores ($r = .54$). Whereas statewide assessment scoring assigns numerical values to student writing, prominent feature analysis or scoring derives numerical values from specific rhetorical features. These results may be helpful for classroom teachers for the assessment and diagnosis of student writing and for professionals who lead staff development programs for teachers.

The purpose of the present study is to identify the prominent features in the writing of 464 seventh-graders, all of whom participated in a statewide writing assessment, and to determine whether or not certain features are associated with state assessment scores, as well as with each other. They all wrote on the same topic, wrote on the same day, wrote under the same time constraints. All the writing was scored through a state assessment process. So there were similar conditions, to be sure. Still their writing was individual, sometimes idiosyncratic, sometimes interesting and sometimes not, but often challenging for the readers to comprehend, not just for what the writing itself says but for what it means ultimately for instructional purposes.

The study was sponsored by the National Writing Project and focused on students in schools in which professional development in writing had been conducted. Staff developers and administrators in the participating schools have asked, "What writing strategies are students using?" and, most important, "What writing strategies and features of writing correlate to success on the state writing assessment?" The present study grew from interest in these questions.

Background

Assessment and Instruction.

In his foreword to Hillocks' *The Testing Trap*, Miles Myers affirms what teachers have commonly known for a long time—that it is state writing assessments, rather than state standards, that drive writing instruction (Hillocks, 2002). Hillocks interviewed over 390 practicing educators in five states (Illinois, Kentucky, New York, Oregon, Texas) on "theories of writing underlying the assessments, the kinds of writing tested, the scoring criteria, scoring procedures, materials for teachers, and a variety of dimensions of teacher-reported practices" (p.19). Results from each state showed a direct influence of the characteristics of the writing assessment on writing instruction. The main rhetorical focus of teachers in four of the five states was current traditional rhetoric, followed by expressivist rhetoric, epistemic rhetoric, grammar, and writing about literature. The exception was New York state, where the main focus was writing about literature (p. 191). The secondary focus in all five states was grammar (p. 192). Hillocks further contends that if the quality of teaching is to be improved, then intervention must come at the teaching level, rather than the assessment level. Yet teachers in the five states in the study, while finding fault with their state writing assessments, had neither the training nor the inclination to conduct detailed analyses of their state assessments. Instead, teachers by and large adjusted their teaching to fit the assessments.

In studying the effects of state assessments on writing instruction, Huot (1996) argues for a new theory of assessment, one that will "recognize the importance of context, rhetoric, teaching, and learning" (p. 552) and be based on clear features of language use. Huot calls for a system that will assess writing within the context in which it was written and on its own merits rather than compare it to some numerically based guideline, so that what is assessed is also what is taught. While Huot advocates the use of qualitative procedures (interviews, observations, thick description) in order to create assessment that is sensitive to a specific context, he also asserts that evaluation must do more than judge a piece of writing; assessment must "be able to describe the promise and limitations of a writer working within a particular rhetorical and linguistic context" (p. 564).

The authors have observed that teachers spend considerable time studying the returned state-scored papers, looking for "what the scorers are really looking for" or "the rubric underneath the rubric." Indeed, teachers have stated that the raters must surely be bringing their own criteria to the scoring task. In the process of reviewing returned scored papers, some teachers have put forth hypotheses based on limited anecdotal evidence: "The key is in organization—the five paragraph theme." or "A conclusion should be a restatement of the opening sentence." Some teachers maintain that even though voice is not mentioned on any of the state writing rubrics, the highest scores are generally awarded to papers that exhibit strong voice. The concern here, however, is that some of the ad hoc conclusions formulated by teachers may be false, and that some conclusions may actually prevent improvement in writing achievement.

Studies of holistic raters bear out the assumption that raters are influenced by features of the writing beyond those described by the

rubric. According to Huot (1990), early studies of rater bias show that usage and mechanics rank high with raters. Nold and Freedman (1977) found correlation between raters' scores and the presence of final free modifiers (modifiers occurring after the main clause). Freedman (1979) also found that raters are most concerned with content and organization, and later (Freedman 1987) found that the text itself is a dominant influence on raters. Huot (1990) asserts that these and other studies raise the question as to whether or not it is even possible for a holistic rater to see just those parts of a student's writing that are relevant to a particular rubric. In contrast, the present study focuses on the prominent features observed by readers, how those features correlate with state assessment scores and how they correlate with each other.

Prominent Feature Analysis.

Rather than rereading the returned state-scored papers and comparing them once again to the state rubric as classroom teachers might, we developed a method of analysis that focused on the student writing itself. In examining this large sample of papers, we used no rubrics or guidelines for evaluating writing. Instead, we developed a process in which a team of professional English language arts teachers came face to face with the seventh-grade writing and then recorded what they observed, answering the question, "What stands out as prominent in this paper?" The term "prominent feature analysis" was coined after the fact, not prior to the study. During the discussion prior to the reading of the papers, the leader asked team members to identify, among other features, cumulative sentences and final free modifiers, voice, and certain intersentential connections.

Final Free Modifiers.

The importance of the final free modifier was described by Francis Christensen, who advocated the use of a short base followed by modifiers (Christensen, 1963). He asserted that the form of the sentence itself led writers to generate ideas related to the base clause. He identified five kinds of nonclausal free modifiers: noun cluster, verb cluster, nominative absolute, adjective, preposition. In a succinct review of Christensen's work, Connors (2000) described its strengths and limitations, its brief history, and its unsuccessful pedagogy ("an expensive boxed set of overhead transparencies and workbooks") and the few research projects exploring its effectiveness. Nevertheless, after all these years, the efficacy of the final free modifier continues to show up (Brooks, 1975; Gebhard, 1978; Huot, 1990; Nold & Freedman, 1977). Deserving special mention is "a full-scale empirical research test" conducted by Faigley (1979), who tested four experimental sections of freshmen composition students who were taught the Christensen rhetoric and four control sections who were taught from McCrimmon's *Writing with a Purpose*. The writing of the students who studied the Christensen rhetoric "not only was measurably more mature but also received better average ratings . . . from blind holistic readings" (Connors, 2000, p. 100).

Voice.

Sperling and Freedman (2001, p. 375), influenced by the work of Bakhtin (1986), Casden (1993), and Sperling (1995, 1998), present voice as socially and culturally embedded in both the writer and the reader, putting forth the idea that "students as writers participate in a 'role complex,' their voices reflecting the multiple stances they assume to multiple others both inside and outside of school." Roid's (1994) cluster analysis of direct writing assessments of third- and eighth-grade students in Oregon found that the six trait analytic scores (ideas, organization, voice, sentence fluency, word choice, and conventions) yielded 11 different scoring patterns or cluster types. Whereas about 20% of student papers scored high in all traits at each grade level and another 20% all low, other patterns in the remaining 60% included four patterns that involved voice: 1) high scores in voice along with high scores in ideas and organization, as Roid explained, "the creative or stylistic aspects of writing" (p. 169); 2) low scores in these three traits; 3) scores in the medium range on all traits except for a low score in voice and; 4) scores in the medium or low range on all traits except for a high score in voice.

Others, including Elbow (1994) and Palacas (1989), have theorized about voice, its definition, its role in various modes of writing, and the possibilities for teaching it. Palacas argues that the primary issues concerning the concept of voice are whether it can be objectified, whether it is to remain a question of emotion, and whether it can be taught. Applying a linguistic method, he found one aspect of voice in the parenthetical expression, as a "parallel conversation" between the writer and reader, and went on to propose a "grammar of voice" (p. 128). Responding to the debate as to whether text should be disembodied or whether it should represent a real person, Elbow (1994) suggests that we read with two lenses, one a text lens, reading for what is said, and the other a voice lens, reading for how it is said, and that both are needed to understand fully the written language. Ultimately, the objectification of voice into a teachable curricular element has yet to be realized.

Flawed Sentences.

All syntactically flawed sentences were categorized as having a "weak structural core" (Krishna, 1975). In essence, these are sentences in which the main idea of the sentence does not appear as the main subject-verb but in some peripheral place. "This habitual wasting of the subject-verb position, along with the frantic struggle to fit a central thought into a peripheral expression . . . is the source of many, perhaps most, of the structural errors that appear in student papers, and, I believe, contributes to idiomatic, stylistic, and grammatical errors as well" (Krishna, 1975, p. 45). Fragment sentences were also classified as having a weak structural core. A run-on sentence, however, was classified as a punctuation error, since it does not exhibit the more serious syntactic flaw.

Intersentential Connections.

Several prominent features deal with the whole piece of writing (e.g., repetition, effective organization, coherence/cohesion) and, in some instances, focus on intersentential connections. The rationale for these features is derived from selected rhetorical studies, including Christensen (1965) who identified three approaches to paragraph analysis—coordinate sequence, subordinate sequence, and mixed sequence; Gray and Benson (1982), who translated Christensen's approaches for the classroom; Becker (1965), who identified two paragraph patterns (topic-restriction-illustration and problem-solution) and two basic paragraph markers (equivalence classes and function words); and Witte and Faigley (1981), who described the essential differences between cohesion and coherence in prose. Perhaps the most useful approach in the middle school to intersentential connections, as well as internal harmony within the sentence, is derived from Corbett (1991), who has described the varieties and forms of repetition in the English sentence, much of which can be applied to the paragraph and longer stretches of prose.

Participants and the Writing Assessment

Participants.

Two school districts were selected for this study, two middle schools in District A and one in District B, three schools in all. Both districts have provided extensive staff development programs on writing for their teachers, and both agreed to provide copies of students' writing from the seventh-grade assessment.

The two schools in District A include a student population of 237 seventh graders, 52% of whom qualify for free or reduced price lunches. Ten percent of the student population is African American, 90% is Caucasian. In District B, 63% of the 227 seventh-grade students qualify for free or reduced price lunches. Nineteen percent of the student population is African American, 76% is Caucasian; and 5% of the population is Native American. Per pupil expenditures for District A average \$5,510, for District B, \$5,441.

The Seventh-grade Writing Assessment.

The statewide seventh-grade writing assessment occurs in March, and the papers are returned in late July. The papers provided for the present study were from the 2004 school year. Although the specific prompt for the 2004 assessment was not made public, it can be inferred from the hundreds of samples from the assessment: "Write to describe the activities you enjoy doing outside of school." We further inferred from the hundreds of papers that the prompt called for informative writing. The one former state prompt made public also led us to infer that students were asked to write at least three paragraphs; plan their writing; use a checklist for following the prompt; make use of details, effective words, and complete sentences; organize their pieces; and attend to conventions. We did not address the question of content for two reasons: (a) the prompt itself circumscribed the role of content in the student writing, and (b) we did not know specifically what the prompt was. Although there are some limitations for using assessment papers for research, for example, the number of words in each paper may be limited, the advantages of the controlled situation and the available state scores outweigh the limitations.

The assessment is constructed by a large testing company. The papers are read outside the state and given a score of one (lowest) to four (highest), or a score of zero for "off topic." Each paper is read by one reader. Even though the conditions of the assessment were less than ideal, they nevertheless represent constraints under which we all must work, teachers as well as researchers. The distribution of writing assessment scores for the two districts is shown in Table 1.

Table 1: Distribution of Seventh-grade Writing Assessment Scores for District A and District B, 2004

Scores	4	3	2	1	0
District A	18	126	90	1	2
District B	10	109	100	3	5
Totals	28	235	190	4	7

Note. A score of zero indicates that a paper was written off topic. It does not indicate the quality of the writing. $N = 464$

From Table 1 it is apparent that most papers received a score of 3 or 2. Indeed, 91.5% of the papers fall into these two categories, revealing that the assessment is essentially a two-category analysis. Only four of the 464 students in the study received a score of 1, indicating that less than one percent has very serious problems in written communication. The 28 students who scored at level 4 represent six percent of the total population of the study.

Research Questions.

1. What are the prominent features of the writing of seventh-graders on a statewide writing assessment?
2. What correlations exist between the prominent features and statewide writing assessment scores?

3. What correlations exist among the prominent features themselves?

Research Design and Methods.

For the daunting task of reading and analyzing the papers, a research team was chosen, all experienced English language arts teachers, grades 4 through college. The members of the research team were selected by the principal investigator for (1) prior experience, (2) successful background in classroom teaching, (3) strong leadership, (4) participation in and distinguished service to the National Writing Project. In all, the research team was composed of some 12 individuals, including the authors. The research team members, their qualifications and educational backgrounds are shown in Table 2.

Table 2: Educational Background, Teaching Level and Professional Qualifications of Research Team Members

Researcher	Ed.	Grade level	NWP	NBPTS ²	Other
A	B.A.	9-10	Yes		NSC ⁶
B	Ed.S.	4, adm.	Yes	Yes	AED ³ NSC ⁶
C	B.A.	10-11	Yes		NSC ⁶
D	Ed.S.	8	Yes	Yes	NSC ⁶
E	M.Ed.	4	Yes	Yes	AED ³ ; NSC ⁶
F	M.Ed.	7-8	Yes		NSC ⁶
G*	Ph.D.	7-11, college	Yes		
H	Ph.D.	college	Yes		SDE experience; NSC ⁶
I	B.A.	10-12, college	Yes		NSC ⁶
J	Ed.S.	10-12, college	Yes	Yes	NAEP ⁴ ; NSC ⁶
K ¹	Ph.D.	1, college, adm.	Yes		AWC ⁵ ; NSC ⁶
L	B.A.	10-11	Yes	Yes	NSC ⁶

Note. * Author of multiple volumes on rhetoric and composition. ¹Principle investigator. ²NBPTS refers to National Board certification. ³AED indicates experience in the NWP study of 3rd and 4th grade writing. ⁴NAEP indicates membership on a committee for the National Assessment of Educational Progress. ⁵AWC indicates major development role for the NWP Analytic

The Analysis.

The first meeting of the research team began with a training session. The identification of prominent features depended, not on an established list of features, but on the professional experience of team members. Using identical sets of sample papers, team members discussed what constitutes a level of prominence for each of the features they were observing. The team agreed on the definition of prominent as "distinguishing, not ordinary." Team members looked for tell-tale signs in single words, for patterns of occurrence, for evidence of what might rise to the level of prominence. For example, a single misspelled word might not be worthy of note as faulty spelling; a single metaphor, however, would be worthy of note. The training session sought to establish consistency among the readers. Prominent features are either *there* or *not there*, thus creating a dichotomy. During the actual reading process, team members were encouraged to ask another reader to confirm a problematic decision about the prominence of a feature in a paper. Sometimes a reader would read aloud an interesting or troublesome passage and ask the whole group for an opinion. As the reading of the papers progressed, the leaders observed that the research team developed a spirit of camaraderie and confidence in their ability to identify the distinguishing characteristics of the writing. When the research team concluded the reading, all the papers were then reread by two of the authors to confirm the accuracy and consistency in identifying the prominent features.

During the analysis several important realizations occurred: (a) All the prominent features could be classified as either positive or negative. There were no neutral observations. (b) Some features, such as repetition, could be either negative (as in redundancy) or positive (as in effective anaphora, e.g., Martin Luther King's "I Have a Dream"). (c) Some features were negative (faulty spelling) without a positive corollary. (d) Some features were positive (the presence of a cumulative sentence) without a negative corollary. Team members noticed, however, some overlapping and inconsistency among themselves. For example, some recorded "metaphor," "simile," and "personification" separately while others used the term "metaphor" to characterize all three. It was agreed that the simple term "metaphor" was adequate to describe all kinds of metaphoric language. The team continuously reviewed the language of the descriptors, further refining the definitions. A final list of 32 prominent features emerged, 22 positive and 10 negative. The list of prominent features for the present study, along with definitions and examples, is shown in Table 3.

Table 3: The Prominent Features: Definitions and Examples

Elaborated details--use of vivid, appropriate, or striking details; goes beyond a listing of details

I like going fishing . . . because I like to set back and let the breezy wind roll right through me, and whenever I get a bite and I'm not asleep, I jump up, grab my pole and reel it in. (399)

Sensory Language--language addressing the six senses, including direct quotations

I like softball. I like to hear the crack of the bat on the ball. I love the smell of the fresh cut field. The sting of the ball when it hits me. (126)

Metaphors--all types of metaphoric language (metaphor, simile, etc.); especially noted is the use of common words used in metaphoric ways

When I am out on the skateboard, I am as fast as a bullet. (14)

Swimming is so much fun, as fun as a new puppy wanting to play. (3)

Alliteration--effective repetition of sound in successive words

We rode horses on a ranch where the green grass grew and the fresh water flowed. (327)

Vivid Verbs/Nouns--striking diction, very appropriate and descriptive

By playing with him, I taught him how to sit and fetch. (4)

Hyperbole--exaggeration *I travel to tons of places all over Mississippi, but my favorite place . . . (300)*

Striking Words--striking word usage, including appropriate or surprising nouns, verbs, adjectives, adverbs, etc.

I am a very athletic and versatile young man. . . . I can virtually play anytime I want to. I usually play golf after school with buddies who all live on the golf course. (248)

Usage Problems (-)--occurrences of nonstandard, social, regional, or ethnic dialect features

This are all the reasons why I like do them. I do all them because of every reason I put down on this paper. This are the activities that . . . (442)

Cumulative Sentence--sentence with a base clause and one or more free modifiers

I like the feeling of being strapped down to a 1000 pound muscle machine, twisting and turning and tilting you with his every

move. (401)

Verb Cluster--type of free modifier (-ing or -ed participle)

It is very thrilling to be out in the middle of a lake, reeling in a whopping five pound catfish. (391)

Noun Cluster--type of free modifier; a noun, possibly with attachments

I love playing video games, the excitement to not know what's going to happen next in the game, the long fun that I have while playing the game. (201)

Absolute--type of free modifier; an independent noun with its own verb and deleted auxiliary verb

Hunting shows you nature because of the things you see, like the squirrels going up the tree, speeding by you, the birds chirping and singing their lovely songs, the leaves falling like feathers gently to the ground . . . (262)

Adverbial Lead--beginning the sentence with adverbial (word, phrase, or clause)

Every Wednesday at 5:45 pm, I go to choir practice. Then at 7:00 pm, the choir comes out to sing. When I come out, I feel like a professional singer. (301)

Balance/parallelism--all types of parallel structure

Climbing a steep hill and reaching the top makes me feel like I've conquered the world. (394)

Repetition--repeating the same word, or a form of it, effectively (also includes repetition of phrases)

I know that everyone can visualize me with computers, but my classmates seem to have a problem visualizing me fishing. (391)

Sentence Variety--effective use of a variety of sentence forms and lengths

One activity I enjoy while I am not in school is playing my Playstation 2. I enjoy this because it lets me feel like I'm in the game, shooting, racing, or playing a card game. It also is the only thing I can do, so I guess I'm addicted. To describe it, it is black with the PS2 sign on it. (382)

Weak Structural Core (-)--sentences that are "derailed" with misplaced awkward elements; also includes sentence fragments

I enjoy doing when I'm not at school.

I play basketball to keep me busy from anti-drugs.

I like playing this sport to be able to be good and quick. (265)

Garble (-)--unintelligible sentences

Then I to wele when I heard that name. (522)

Effective organization--clear pattern of organization

• *Splash and giddy-up! My favorite . . . swimming and riding my horses. . .*

• *The warm Southern breeze heats up your face on a sunny summer afternoon . . . we all enjoy the beautiful blue clear water. . .*

• *Riding horses . . . the wind brushing through your hair while you're galloping through the green grassy field still wet with dew makes me feel so alive . . . everyone should be able to enjoy the comfort of swimming and riding horses. . . (283)*

Subordinate Sequence--an organizational pattern that follows a "detail of a detail of a detail" pattern

I always have a blast, well at least when my team wins. It isn't too much fun when we lose and my older brother and his team wins.

I guess it isn't who wins or loses, it's who has the most fun. At least that's what my mom says. (396)

Transitions--the presence of key function words or phrases to enhance organization

Another sport that my friends and I play is baseball. We love baseball so much that every Sunday after church my team and I go to our field and play as hard as we can until it's time to go back to church . . . Everyday after school I go home and play basketball until I can't see because of the dark. Then I have to turn on the court lights. (339)

Coherence/Cohesion--obvious presence of cohesive devices to create cohesion or coherence

I talk on the phone. Everyday after school I call someone to . . . miss out on any gossip . . . My mom complains that I talk on the phone too much, but I could never get enough of chatting . . . I run the phone bill up . . . I love to talk. When I call . . . prepare themselves to be on the phone for about four hours. (311)

Weak organization (-)--obvious lack of organization

(From a paper in which almost every sentence introduces a new topic.)

My favorite activity is to run, play and watch tv sports and to go to school and to do good. I love to spend time with my family and go to church . . . (472)

Redundancy (-)--repeating the same idea or concept over and over; sometimes described as "verbiage" or "mindless filler"
I enjoy a lot of amazing activities That is only one activity I enjoy. . . . Another activity I enjoy is Another amazing activity I enjoy is These are not all of the activities that I enjoy (495)

List Technique-- list of ideas related to a topic but not to each other; undeveloped writing
I like listening to music. I enjoy dancing and singing. I enjoy writing. I like watching TV. I like playing basketball . . . (471)

Voice--the presence of an original, personal or authentic conception of the subject
I love to read! When I cuddle up with a book, I'm in a whole new world. I can go places I've never been to before. Who would've thought that pieces of paper would be so extraordinary? (315)

Addresses Reader--direct communication with the audience
Come with me, and I will tell about the activities I enjoy. . . (150)

Narrative storytelling--including stories, anecdotes, and personal experiences to develop ideas
(From a paper about playing football)
My favorite NFL team is the Green Bay Packers. They won eleven games this year and lost five. They went to the playoffs and won the wildcard (483)

Faulty Punctuation (-)--patterns of any/all varieties of punctuation errors
The activity I enjoy doing when I am not at school. Is to pretend that I am acting to pretend that. I am a famous actor starring in a famous movie. . . . Another thing I like to do as an out of school activity. Is to read read everyday when I get home. (174)

Faulty Spelling (-)--a pattern of faulty spelling
I sometimes go to the church after school dearing the week . . . I very often go fishing dear in the week. I have studying to do most of the time. . . I go on vacations with my parents . . . That's mostly what I don when I am out for along period of time. (496)

Shifting Point of View (-)--abrupt changes in the writer's point of view or subject
(Final paragraph of first person paper)
Now that I have told you about my activities I enjoy. I hope it has put much enthusiasm in your heart to go try the same things. . . . Well that's the end of my informative paper. (466)

Illegible Handwriting (-)--handwriting so faulty it is almost unreadable

Note. Numbers in parentheses represent the student paper number. Samples are in italics. Negative features are indicated by (-).

Classification Consistency.

Although interrater reliability, commonly associated with holistic scoring, cannot be applied to prominent feature analysis, it is possible to establish a level of classification consistency among team members involved in the analysis. Agreement in prominent feature analysis can be expressed as a measure of classification consistency, with the identification of prominent features treated as a form of classification.

To establish the level of classification consistency achieved in the prominent feature analysis, the authors examined the individual score sheets for each of the 464 papers, determining how many changes were made in prominent features assigned from the initial analysis of features through the final reading. There were 484 changes in the prominent features assigned to the entire set of papers across the multiple readings. There was a possibility of 14,848 changes, considering that there were 32 features, and that each of the features originally assigned to the writing could have been deleted and each feature not assigned could have been added. The percentage of agreement in this case is 97%. In other words, the data resulted from a consistent process of assignment or classification which was then further refined. The judgments of presence or absence of prominent features are therefore considered to be both highly consistent across independent readers and to have yielded credible data for the analyses reported herein.

Developing the Prominent Feature Score.

Observations revealed by the analysis led to the development of the prominent feature score. The team observed that prominent features may be either positive or negative. Not a single instance of a neutral prominent feature was found. In all, the team identified 32 prominent features, 22 positive aspects, such as the presence of vivid verbs or nouns, and 10 negative aspects, such as faulty punctuation. Within a given piece of writing there is a "still point" between the positive and the negative. A piece of writing with no prominent features would reside at this still point. A paper with an equal number of positive and negative features would also reside

at this point. From these observations, the team developed the prominent feature score through the following steps.

- (1) The still point was assigned the value of 0.
- (2) Each prominent feature was assigned the value of +1 or -1. There was no weighting of individual features relative to each other.
- (3) The features were then summed, yielding a scale of -10 to +22.
- (4) In order to report the prominent feature score as a positive number, 10 points were added to each score. Thus, the still point became 10 instead of 0.

Theoretically, prominent feature scores in the present study may thus range from 0 to 32. In actuality the observed range of prominent feature scores was from 3 to 21.

Correlations and Statistical Analyses

Detailed analyses of the 32 aspects were completed for each of 464 papers. Table 4 gives summary statistics and correlations of the summed prominent feature scores (positive, negative and total) with the state writing score. As expected, the summed positive features and the total prominent feature scores are positively correlated with the state writing score, whereas the summed negative feature score is negatively correlated with the state writing score. If cases having a state writing score of zero—meaning an off-topic paper or no serious attempt at an answer (seven in all)—are omitted, the correlations of the prominent feature scores are stronger by .02 to .04. Interestingly, the mean score for the prominent features ($M = 10.3$, where M is mean) corresponds to the still point of 10 present in the prominent feature score formula.

Table 4: Correlations of Prominent Feature Scores and State Writing Scores

Measure	Correlation to State writing score	Mean	SD
Prominent feature score	.54*	10.30	3.19
Positive features	.48*	2.28	2.20
Negative features	-.48*	1.97	1.40

Note. Only cases having state writing scores greater than zero are included. $N = 464$. Correlations of $|.12|$ or more (*) are statistically significantly different from zero at the .01 level.

Distribution of Prominent Feature Scores.

Table 5 reveals the distribution of prominent feature scores, which range from 3 to 21. This distribution generally follows a normal curve, the majority of the scores occurring between 7 and 14.

Table 5: Frequency Distribution of Prominent Feature Scores

Score	Frequency	Percent	Cumulative Percent
3	2	0.4	0.4
4	2	0.4	0.9
5	9	1.9	2.8
6	30	6.5	9.3
7	53	11.4	20.7
8	67	14.4	35.1
9	47	10.1	45.3
10	49	10.6	55.8
11	49	10.6	66.4
12	33	7.1	73.5
13	43	9.3	82.8
14	31	6.7	89.4
15	20	4.3	93.8
16	15	3.2	97.0
17	6	1.3	98.3
18	3	0.6	98.9

19	2	0.4	99.4
20	2	0.4	99.8
21	1	0.2	100.0
Total	464		

Comparisons by State Writing Score Levels

Table 6 summarizes the percentage of papers at each score level (0, 1, 2, 3, 4) that exhibited discrete prominent features. The elements, such as sensory language, were simply scored as present or absent. The general trend is for positively scored elements to have higher rates of occurrence with successively higher state writing scores, whereas negatively scored elements, such as weak structural core, have lower rates of occurrence with successively higher state writing scores. Cases having state writing scores of zero did not necessarily follow either pattern, though there were very few such papers in the batch ($n = 7$). Similarly, there were only four cases having state writing scores of 1, so some of the relative percentages at this score level may not correspond to the patterns described.

Table 6: Percentages of Papers that Exhibited Prominent Feature Elements by State Writing Score

Feature	State score (%)				
		0	1	2	3
Elaborated details	14	25	15	37	68
Sensory language	14	0	8	15	43
Metaphor*	14	0	9	26	36
Alliteration	0	0	0	1	0
Vivid verb/noun	0	0	3	7	32
Hyperbole	0	0	0	0	7
Striking words	0	0	9	18	50
Usage problems (-)	57	50	34	24	4
Cumulative Sentences*	0	0	5	14	18
Verb clusters*	0	0	4	14	18
Noun clusters*	0	0	1	0	4
Absolutes*	0	0	1	1	0
Adverbial leads	0	0	4	7	25
Balance/parallelism	14	0	3	14	32
Repetition	14	0	2	5	18
Sentence variety	0	0	14	26	43
Weak structural core (-)	29	50	29	11	0
Garbles (-)	14	25	2	0	0
Effective organization	0	0	14	30	75
Subordinate sequence	0	0	1	1	0
Transitions	0	0	5	10	25
Coherence/cohesion	0	0	1	9	7
Weak organization (-)	29	4	17	11	0
Redundancy (-)*	43	75	69	47	25
List technique (-)	0	0	12	7	0
Voice	29	25	17	38	43
Address reader	0	0	1	0	0
Narrative storytelling	43	25	8	7	11
Faulty punctuation (-)	14	75	20	13	0
Faulty spelling (-)	6	75	30	23	4

Shifting point of view (-)	43	25	24	22	11
Illegible handwriting (-)	0	50	8	1	0
Extra	0	0	3	3	0
Number of cases	7	4	189	236	28

Note. Elements are positively scored except those indicated by "(-)."

Elements noted with "*" could have multiple instances. $N = 464$.

A score of 0 typically refers to a paper that is considered "off topic" rather than to the quality of writing.

Statistical correlations are shown below. Tables 7 and 8 reveal prominent features that correlate significantly with state scores. Teachers and staff developers might consider this array when planning instruction for student writers.

Table 7: Kendall's tau b Correlations of Positive Features with State Writing Scores

Feature	State Score	
	Coefficient correlation	Significance
Elaborated details	.289*	.000
Sensory language	.183*	.000
Metaphors	.222*	.000
Alliteration	.046	.305
Vivid verbs/nouns	.192*	.000
Hyperbole	.115*	.010
Striking words	.208*	.000
Cumulative sentences	.152*	.000
Verb clusters	.174*	.000
Noun clusters	-.007	.875
Absolutes	.011	.813
Adverbial leads	.144*	.001
Balance/parallelism	.229*	.000
Repetition	.128*	.004
Sentence variety	.112	.013
Effective organization	.278*	.000
Subordinate sequence	.002	.967
Transitions	.154*	.001
Coherence/cohesion	.081	.073
Voice	.214*	.000
Narrative storytelling	.035	.440
Address reader	-.010	.823

Note. Asterisk (*) indicates significance at the .001 level. $N = 464$.

Table 8: Kendall's tau_b Correlations of Negative Features with State Writing Scores

Feature	State Score	
	Coefficient correlation	Significance
Usage problems	-.167*	.000
Weak structural core	-.247*	.000

Weak organization	-.146*	.000
Garbles	-.146*	.011
Redundancy	-.229*	.000
List technique	-.086	.055
Faulty punctuation	-.155*	.001
Faulty spelling	-.172*	.000
Shifting point of view	-.069	.125
Illegible handwriting	-.215*	.000

Note. Asterisk (*) indicates significance at the .001 level. $N = 464$.

Regression Analyses

Two sets of multiple linear regression analyses were run, one using the summed positive feature and summed negative feature scores as the independent variables, and the second using all of the individual prominent feature scores as independent variables. In each instance, the state writing score was used as the dependent variable. Cases having a score of zero (off topic) were omitted from these analyses, as the writing may well have had noteworthy characteristics even though no credit was awarded based on the state writing score rubric. This omission yielded 456 cases.

For the first analysis, forcing both independent variables (total positive features score and total negative features score) into the model yielded a statistically significant squared multiple correlation with state writing scores, $R^2 = .29$, $F(2, 453) = 94.22$, $p < .001$. That is, the summed positive features score and summed negative features score could, together, explain 29% of the variance in state writing scores. Standardized regression coefficient estimates were 0.31 for the positive features score and -0.30 for the negative features score. Unstandardized coefficients were 0.09 and -0.14, respectively. This implies that, on average, two students whose positive feature scores were 6 points apart would have estimated state writing scores slightly more than 0.5 points apart, when negative feature scores are held constant. Likewise, two students whose negative feature scores were 4 points apart would, on average, have estimated state writing scores slightly more than 0.5 points different, when positive features scores are held constant. Clearly, both positive and negative features help to explain differences among students' state writing scores.

For the second analysis, a backward elimination method was used such that all of the individual prominent feature elements were initially forced into the model (again, using state writing score as the dependent variable), then any feature score that did not have an estimated regression coefficient significantly different from zero at the .05 level was eliminated, then all coefficients were re-estimated. This continued until all remaining independent variables had statistically significant estimated regression coefficients. Because any step-based method can capitalize on the idiosyncratic nature of the sample, this was done in two stages. In the initial stage, an approximate 2/3 sample ($n = 311$ cases) was randomly selected from the full data set. The results yielded a statistically significant model, $F(10, 300) = 17.17$, $p < .001$ with an overall $R^2 = .36$. The 10 independent variables that were retained included five of the positive feature elements (elaborated details, sensory language, metaphor, hyperbole, and effective organization) and five of the negative feature elements (faulty spelling, weak organization, faulty punctuation, redundancy, and illegible handwriting). This 10-predictor model was then applied to the remaining ($n = 145$) cases as a cross-validation. The regression coefficients, separately estimated for the two subsamples, were observed to correlate very well, $r = .93$, thus making the case that the model performed similarly across the samples. The subsamples were combined and the 10 chosen feature elements were forced in as independent variables to arrive at the final estimates of the regression coefficients and of the overall R^2 , again using state writing scores as the dependent variable. The R^2 for this 10-predictor model, applied to all cases having state writing scores above zero, was .32. To put in context the explanatory power of the summed square of scores for differences of state writing scores, Cohen (1988) gives us guidelines for interpreting correlations in social science research and suggests that squared correlations of .14 to .16 would indicate a large effect size.

Estimated regression coefficients from this 10-predictor model using state writing scores as the dependent variable are given in Table 9. In each case, the sign of the regression coefficient is congruent with the type of feature. That is, positive feature elements have estimated regression coefficients that are positive and negative feature elements have estimated coefficients that are negative. Further, with the possible exception of elaborated details, the standardized regression coefficients are all fairly close in value, suggesting that each of these prominent features contributes comparably to the overall explanatory power of the model. Inspection of the unstandardized regression coefficients might cause the reader to conclude that the positive feature of hyperbole, with an estimated coefficient of .96, is somehow an extraordinarily potent explanatory variable. This value signifies that, for papers in which hyperbole is a prominent feature, one would estimate a state writing score that is higher by nearly a full point, were all other feature values held constant. However, Table 6 indicates that only a small fraction (7%) of all papers had hyperbole as a prominent feature and these all came from maximum-score (state writing score = 4) cases. It would be interesting to ascertain why only the very highest scoring papers were likely to have used instances of hyperbole. As a set, these 10 prominent features served to explain

approximately 32% of the variance in state writing scores for the full data set.

Table 9: Estimated Regression Coefficients for Prominent Feature Elements on State Writing Scores

Estimated Regression Coefficient		
Feature	Unstandardized	Standardized
(Positive features)		
Elaborated details	0.28	0.21
Sensory language	0.22	0.12
Metaphor	0.14	0.15
Hyperbole	0.96	0.10
Effective organization	0.20	0.14
(Negative features)		
Weak organization	-0.22	-0.14
Redundancy	-0.16	-0.13
Faulty punctuation	-0.23	-0.14
Faulty spelling	-0.14	-0.10
Illegible handwriting	-0.31	-0.10
Constant	2.64	

(N=456)

Both sets of analyses, the first using summed positive and negative feature scores and the second using individual feature elements, clearly show that both types of elements are important in accounting for differences in the state writing scores. However, since more than two-thirds of the variance in state writing score is unexplained by the prominent feature scores, other characteristics of the essay must be influential in the holistic scoring of the writing samples. Two technical concerns about the criterion variable of state writing scores must be mentioned here. First, one consideration is the somewhat restricted range of score values; though the state scoring rubric offers potential scores of 1-4 (and 0 for off-topic or no credible response), nearly 92% of the observed scores were 2 or 3. Hence, there is effectively a restriction of range, which most likely will serve to depress the strength of any correlations observed. A second concern is the fact that the state writing scores might themselves not be highly reliable. Hence, some of the variance in state writing scores that is not accounted for by the prominent features scores may well be unreliable variance. Given the current data set, there is no definitive way to ascertain to what extent the unexplained variance in state writing scores may be a function of (a) restriction of range of assigned scores; (b) unreliability; (c) other systematic aspects (e.g., scorer effect); or some combination of these factors.

Intercorrelations of Prominent Feature Elements

Each of the feature elements, scored as 1 if present and 0 if absent, was correlated with all of the others. Correlations among the prominent feature element scores are given in Table 10. Of the individual elements, voice shows the most consistent, positive correlations with others. The strongest single correlation observed was that between cumulative sentence and verb cluster ($r = .83$). Of the 496 correlations among the prominent features elements, 108 (22%) were statistically significantly different from zero at the .001 level. However, many of these significant correlations were observed to be low in absolute value, less than + .20. One possible reason is that several of the elements were not frequently observed (see Table 6). Were we to have had the luxury of combining prominent feature element appearances for students across multiple writing samples, the correlations may well have been different. In any event, the overall strength of many of these relationships in the present data set was low. To the extent that these low correlations are an accurate portrayal of relationships among the prominent feature elements in the population, one could conclude that there is considerable uniqueness among the writing elements.

Table 10: Intercorrelations of Individual Prominent Feature Elements

Table 10:<i>Intercorrelations of Individual Prominent Feature Elements</i>

Table 10 continued:<i>Intercorrelations of Individual Prominent Feature Elements</i>

Results and Findings

This section is organized into two parts, Positive Prominent Features and Negative Prominent Features. Since regression analyses and correlations among individual features were presented earlier, this section presents major trends and tendencies.

Positive Prominent Features.

For the sake of clarity and economy, we have deleted infrequently occurring positive features and have organized the remaining features into ten categories. Nevertheless, some overlapping among features still occurs, primarily because of the fluid nature of language itself.

Diction. Included here are those prominent features that reflect appropriate word choice: sensory language, vivid verbs and nouns, striking words, and transitions. Table 6 shows that in each of these features, Level 4 writers are clearly superior. Table 7 reveals that each of these features has a significant correlation ($p < .001$) with scores on the state assessment. When these features are combined, the 193 papers at levels 1 and 2 exhibit 49 instances of good diction, whereas the combined 264 papers at levels 3 and 4 reveal 146 such instances, roughly three times as many.

Elaboration. Sixty-eight percent of Level 4 writers and 37% of Level 3 writers used elaborated detail. Only 16% of the combined level 1 and 2 papers displayed this feature. Beyond its inherently positive quality, the use of elaborated details serves another important function, namely, to reduce the problem of redundancy. Those writers who elaborated on a topic were less likely to repeat the same ideas over and over. This is especially borne out among the District B Level 4 writers. Of the 9 students who scored at Level 4, six exhibited the use of elaborated details, and none of those 6 papers exhibited redundancy. The one instance of

redundancy among the Level 4 papers was in a paper that did not exhibit elaborated details. The same trend is present in District A, though not as pronounced.

Adverbial leads. This category includes single adverbs, adverb phrases or adverb clauses used at the beginning of a sentence. In the seventh-grade assessment, adverbial leads appear most prominently at Level 4, where 25% of the students used them. At Level 3, only 7% used them; at Level 2, 4%; at Level 1, none. Adverbial leads correlate positively ($p < .001$) with the cumulative sentence and verb cluster, which is interesting since emphasis in the cumulative sentence has been on *final* free modifiers. Adverbial leads also correlate positively ($p < .001$) with voice.

Metaphor. The single term *metaphor* accounts for all types of symbolic language, including simile, personification, and others. Frequency of occurrence of metaphor is positively associated ($p < .001$) with quality of writing on the assessment. At Level 4, 36% of students employed metaphor; at Level 3, 26%; at levels 2 and 1 combined, 9%. Metaphors exhibit a statistically significant correlation ($p < .001$) with vivid verbs, vivid nouns and voice. It appears that the effective use of metaphor is an important if not crucial aspect of the writing curriculum as it relates to raters' perception of writing quality.

Repetition, balance and parallelism. Two features are combined in this category, repetition and balance/parallelism. The term *repetition* here means "good repetition," not redundancy. Table 6 shows that both repetition and balance/parallelism are associated with higher levels of writing, repetition occurring in 18% of Level 4 papers and 5% of Level 3 papers; balance and parallelism in 32% of Level 4 papers and 14% of Level 3 papers. Both features exhibit a significantly positive correlation ($p < .001$) with state assessment scores.

The cumulative sentence and final free modifiers. Four features are included in this category, the cumulative sentence, verb cluster, noun cluster, and absolute. Table 6 shows that each of these features is characteristic of writing at scoring levels 3 and 4. Table 7 shows that the cumulative sentence and verb cluster exhibit a significantly positive correlation ($p < .001$) with state assessment scores. Furthermore, Table 10 shows that a positive correlation ($p < .001$) occurs between the cumulative sentence and metaphors, adverbial lead, vivid noun and verb, balance and parallelism, effective organization, and voice, supporting what others have surmised, namely, that learning the form of the cumulative sentence leads toward tacit understanding of other rhetorical features as well (Faigley, 1979).

Sentence variety. Although the term sentence variety suggests a combination of simple, compound, and complex sentences, the research team agreed that sentence variety is essentially the modulation of sentence length, which gives rhythm and "music" to writing. Table 6 shows that sentence variety occurred in 43% of Level 4 papers, 26% of Level 3 papers, and 14% of Level 2 papers. Table 7 shows that although sentence variety correlates positively with assessment scores, the correlation is not statistically significant.

Intersentential correlations. The item of analysis most prominent here is "well organized." Two related features, subordinate sequence and coherence/cohesion, are not discussed because their frequency of occurrence was too low. Effective organization correlates significantly ($p < .001$) with the state assessment scores as well as with several other positive prominent features. This cluster of correlations suggests the possible presence of a kind of holistic grounding in rhetorical ability.

Narrative storytelling. Table 7 shows no significant correlation between the use of narrative and state assessment scores; the presence of narrative is distributed rather evenly among the scoring levels. Table 6 shows that 11% of papers at Level 4 employed narrative; 7% at Level 3; and 8% at Level 2. Of the eleven papers at levels 1 and 0, four employed the narrative technique, roughly 36%. None of the other prominent features exhibited this widespread distribution among the scoring levels.

Voice. Using the definition of voice as an "original, personal, or authentic conception of the subject" (National Regional Educational Laboratory, n.d.), the research team identified a total of 137 of the 464 papers (30%) as having voice. Voice has a significantly positive correlation ($p < .001$) with the state assessment scores. Although this feature is primarily associated with the upper levels of assessment (30% of Levels 3 and 4 papers combined included voice), the lower levels, 2, 1, and 0, were not totally devoid of voice. Thirty-six of the 200 papers at these lower levels (18%) exhibited voice. These results are consistent with those of Roid's cluster analysis (1994), which revealed that papers at all levels exhibit voice. Although voice is strongly correlated to higher scores, its presence is not limited to a particular score level.

Voice is more abstract than most of the other prominent features; and yet it correlates significantly ($p < .001$) with several concrete features, namely, elaborated details, sensory language, metaphors, striking words, cumulative sentence, verb cluster, absolute, adverbial lead, balance/parallelism, repetition, sentence variety, effective organization, transitions, and coherence/cohesion. That voice yielded more statistically significant intercorrelations than did any of the other features suggests that voice may indeed be taught, or perhaps, may emerge, through teaching these more concrete rhetorical forms.

Negative Prominent Features.

Even though the overall emphasis in the composition curriculum is on positive aspects of writing, some attention must be given to

the function of error. In this section the focus is on those places where seventh-grade writing derailed.

Redundancy. The problem of redundancy proved to be a major negative feature in the seventh-grade assessment. Of the total 464 students in the study, 254 (55%) exhibited redundancy in their writing. Although redundancy is generally associated with the lower levels of assessment, it occurs at higher levels as well. At Level 4, 25% of the papers were characterized as redundant, at Level 3, 47%; Level 2, 69%; Level 1, 75%. The relationship between redundancy and state scores has a significantly negative correlation ($p < .001$). Redundancy also correlates negatively with the presence of the cumulative sentence, sensory language, metaphors, vivid nouns and verbs, striking words, effective organization, and voice.

The presence of redundancy may be the result of the error-avoidance strategies promoted by traditional composition curricula. If someone repeats the same idea over and over, students may reason, the possibility of making an error is reduced. Strategies for reducing redundancy may lie in stressing the positive features named above that correlate negatively.

Weak organization and the list technique. The lack of organization is clearly associated with the lower levels of assessment. There were no instances of poor organization at Level 4, but 27 instances (11%) at Level 3. At Level 2 the frequency of occurrence was 51 (17%). All four students at Level 1 exhibited poor organization. Weak organization exhibits a significantly negative correlation voice and several other positive features as well as state scores.

Weak structural core and garbles. Sentences with a weak structural core follow a definite pattern. There were no occurrences among Level 4 papers. At Level 3, 11% exhibited this feature, perhaps preventing them from reaching the top level. At Level 2, 29% of the papers, almost one out of every three, exhibited such sentences. Although the number of papers at Level 1 was too small for analysis, two exhibited sentences classified as "weak structural core."

The negative correlation ($p < .001$) existing between the presence of a weak structural core and effective organization suggests a failure at both the sentence level and the paragraph level. Young people who do not effectively organize discourse also exhibit problems in organizing their sentences. Weak structural core correlates negatively ($p < .001$) with state scores and positively ($p < .001$) with other negative features. Only six examples of garbles (unintelligible expressions) occurred, four at Level 2 and one each at levels 1 and 0.

Usage. As one would predict, usage problems occurred primarily at the lower levels of assessment: at Level 2, 34%; Level 3, 24%; but at the highest level, Level 4, only 4%. Problematic usage exhibits a significantly negative correlation ($p < .001$) with state scores as well as several positive features.

Shifting point of view. Table 6 indicates that roughly one of every five papers on the writing assessment exhibited problems in shifting point of view. Such shifts correlate negatively with assessment scores, but the correlation is not statistically significant, indicating that shifting point of view is a legitimate concern across the spectrum of writing abilities.

Some teachers, aware that the state rubric penalizes students for a shifting point of view, have informally expressed the belief that the presence of the word *you* indicates a shift in point of view, and have therefore told their students not to use *you* in their writing. The research team informally looked for this occurrence and found the belief to be false. Several papers at Level 4 included the word *you*. A shifting point of view indicates more serious problems, such as a disjointed organizational plan or a writing voice going in several directions.

Faulty punctuation and faulty spelling. The research team looked for patterns of problems, rather than isolated incidents, in punctuation and spelling. As one would predict, both have statistically significant negative correlations ($p < .001$) with state assessment scores. At Level 4, three of the 28 papers exhibited a pattern of misspelling, but none exhibited problems in punctuation. At Levels 1, 2, and 3, roughly one paper in four presented such problems.

Faulty spelling and faulty punctuation both exhibit a statistically significant negative correlation ($p < .001$) with effective organization and a statistically significant positive correlation ($p < .001$) with weak structural core. This suggests that both punctuation and spelling errors may improve as students deepen their understanding of organization and the complexities of syntax. Further study of this kind of tacit learning is needed.

Illegible handwriting. Whereas most activities in composition are cognitive, handwriting is essentially a motor skill and may be beyond the ability of many classroom teachers to improve. Illegible handwriting exhibits a significantly negative statistical correlation ($p < .001$) with state assessment scores, but we have no way of knowing to what extent the scores are influenced by motor ability. In other words, if the handwriting were better, would the score be better? Illegible handwriting exhibits significantly positive correlations ($p < .001$) with weak organization and faulty punctuation. However, there are no significantly negative correlations with the major positive features, suggesting that one cannot judge the quality of writing by handwriting.

Conclusions and Recommendations

Implications for Instruction.

The writing in this study resulted from first-draft, prompt-driven, on-demand writing assessment. Students, relying solely on their internalized repertoire of writing strategies, received no feedback from peers or teachers during the writing task. In Vygotskian terms, analysis of the on-demand writing allowed us to examine the writing strategies students were able to carry out on their own, independently, without adult or peer response (Vygotsky, 1978, p. 210).

The instructional situations in the classrooms of the 464 seventh graders, however, were vastly different from the highly structured assessment situation. Instructional strategies resembled those described by Vygotsky (1978) as in the "zone of proximal development," employing collaboration among learners and between teacher and learners. Students learned to use strategies and stylistic devices in context, through naturalistic approaches in which the elements were woven into the fabric of authentic writing. We have observed that writing instruction in these seventh-grade classrooms included the use of student choice of topics, multiple drafts, peer or teacher response, multiple revision strategies, editing, and publishing, those practices widely accepted as important in the teaching of writing (National Writing Project & Nagin, 2003, p. 92). We have seen evidence that the seventh graders were able to internalize and use a variety of organizational and stylistic approaches that their teachers had introduced. Our findings offer support for teaching key rhetorical elements associated with the overall quality of writing in concert with the conventions that support them, thus helping students to express their ideas effectively and competently. We therefore recommend that teaching strategies for the prominent features identified in this study be carefully designed to be presented and practiced *in the context of student writing* rather than in synthetic or isolated lessons.

Prominent Feature Analysis and Scoring.

Prominent feature analysis is a method of diagnosis; the prominent feature score, derived from the diagnosis, is a method of assessment. The analysis of seventh-grade writing in the present study began with a simple question, "What stands out as prominent in this paper?" The readers were experienced language arts teachers, professionals familiar with the writing of young people. They were able to identify a key element when it appeared subtly, as a single occurrence, for example, metaphor. By contrast, they were able to distinguish between a single misspelled word, which might be a common error or slip-of-the-pen, and a pattern of occurrence that had risen to the level of prominence. As the analysis progressed, team members discussed the emerging elements and came to common understanding. The final result, while not perfect, reflects remarkable concurrence.

The features were identified for the present study, and in other contexts such a list could be and should be refined. Although some features may be included on all such lists, others will come and go with genre, age of writer, and opportunities for learning various writing strategies. Even though the features themselves may vary, the overall method of prominent feature analysis is a valid approach to diagnosis and assessment. The analysis identifies what stands out in the writing, not just simple mistakes or elements of a preconceived formula. It is specific; a feature is either *there* or *not there*. Thus in conferences with students, teachers may call attention to specific elements, both positive and negative.

Prominent feature analysis relies on the education, experience and internal wisdom of the readers, not on rubrics, guidelines, standards or theoretical frameworks from external sources. This permits the readers to focus all attention on the writing itself, without constant reference to outside sources, which may themselves interfere with good evaluation.

Prominent feature analysis is different from other kinds of assessment, such as traditional grading or holistic or analytic scoring, in which numerical values are assigned to abstract rhetorical concepts, such as style, content, or overall quality. In prominent feature analysis, numerical values are *derived* from the presence of specific rhetorical features. In other kinds of assessment, the primary task is to distinguish between numerical values (Is this paper a 3 or a 4, or a B- or C+?). In prominent feature analysis, the primary task is to determine whether or not a rhetorical feature has risen to the level of prominence.

This range of prominent feature scores is obviously greater than the 4-point range of scores provided in the state assessment, and thus appears to be more powerful for discriminating among pieces of writing. Furthermore, beyond the score lies a set of features, both positive and negative, for each paper, verbal descriptors which themselves may serve as diagnostic tools. Patterns that emerge from the overall diagnosis are potentially useful for classroom instruction, program evaluation, and staff development programming. It appears to us that prominent feature scoring solves, or at least represents a first step in solving, the issue of subjectivity in the assessment of writing.

In any profession there is the possibility for misdiagnosis. Specialists in all professions sometimes make mistakes, and so we call for second opinions. So too, in prominent feature analysis mistakes may occur, especially if readers focus on errors or isolated or irrelevant instances. Although prominent feature analysis can obviously be done by a single reader, we recommend that at the outset it be done by a well-qualified team of teachers who engage in professional conversations about the papers at hand. Ultimately, what better resource do we have than professional judgment?

Since the research has been completed, our work in classrooms and in professional development sessions for teachers has confirmed our belief in the value of prominent feature analysis. We hope the larger professional community will explore these ideas

and methods in the pursuit of what we all care about, helping young people express themselves more effectively in writing.

References

- Bakhtin, M. M. (1986). *Speech genres and other late essays*. Austin, TX: University of Texas Press.
- Becker A. L. (1965). A tagmemic approach to paragraph analysis. *College Composition and Communication*, 16, 237-242.
- Brooks, L. (1975). The effects of a study of generative rhetoric on the syntactic fluency of seventh graders. Unpublished doctoral dissertation. Auburn University, Auburn, AL.
- Cazden, C. B. (1993). Vygotsky, Hymes, and Bakhtin. In E. A. Forman, N. Minick, & C. A. Stone (Eds.), *Contexts for learning: Sociocultural dynamics in children's development* (pp. 197-212). New York: Oxford University Press.
- Christensen, F. (1963). A generative rhetoric of the sentence. *College Composition and Communication*, 14, 155-161.
- Christensen, F. (1965). A generative rhetoric of the paragraph. *College Composition and Communication*, 16, 144-156.
- Connors, R. J. (2000). The erasure of the sentence. *College Composition and Communication*, 52, 96-125, 155-161.
- Corbett, E. J. (1991). *Classical rhetoric for the modern student* (3rd ed.). New York: Oxford University Press.
- Elbow, P. (1994). Writing first! *Educational Leadership*, 62(2), 8-13.
- Faigley, L. (1979). The influence of generative rhetoric on the syntactic maturity and writing effectiveness of college freshmen. *Research in the Teaching of English*, 13, 197-206.
- Faigley, L. (1980). Names in search of a concept: Maturity, fluency, complexity, and growth in written syntax. *College Composition and Communication*, 31, 291-300.
- Freedman, S. W. (1979). How characteristics of student essays influence teachers' evaluations. *Journal of Educational Psychology*, 71, 328-338.
- Freedman, S. W. (1987). Summary of the Research. In S. W. Freedman, C. Greenleaf, & M. Sperling (Eds.), *Response to student writing* (NCTE Research Report No. 23, pp. 155-159). Urbana, IL: National Council of Teachers of English.
- Graham, S., & Perin, D. (2007). A meta analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, 99(3), 445-476.
- Gray, J. & Benson, R. (1982). *Sentence and paragraph modeling*. Berkeley, CA: Bay Area Writing Project.
- Hillocks, G., Jr. (1986). *Research on written composition: New directions for teaching*. Urbana, IL: ERIC Clearinghouse.
- Hillocks, G., Jr. (2002). *The testing trap: How state writing assessments control learning*. New York: Teachers College Press.
- Huot, B. (1990). The literature of direct writing assessment: Major concerns and prevailing trends. *Review of Educational Research*, 60(2), 237-263.
- Huot, B. (1996). Toward a new theory of writing assessment. *College Composition and Communication*, 47, 549-566.
- Krishna, V. (1975). The syntax of error. *The Journal of Basic Writing*, 1, 43-49.
- National Writing Project & Nagin, C. (2003). *Because writing matters*. San Francisco: Jossey Bass.
- Nold, E. W., & Freedman, S. W. (1977). An analysis of readers' responses to essays. *Research in the Teaching of English*, 11(2), 164-174.

Northwest Regional Educational Laboratory. (n.d.). *Interactive Six Trait Writing Process*. Retrieved July 7, 2006, from <http://senior.billings.k12.mt.us/6traits/>

Palacas, A. L. (1989). Parentheticals and personal voice. *Written Communication*, 6(4), 506-527.

Roid, G. H. (1994). Patterns of writing skills derived from cluster analysis of direct-writing assessments. *Applied Measurement in Education*, 7(2), 159-170.

Sperling, M. (1995). Uncovering the role of role in writing and learning to write: One day in an inner-city classroom. *Written Communication*, 12(1), 93-133.

Sperling, M. (1998). Teachers as readers of student writing. In N. Nelson & R. Calfee (Eds.), *The reading-writing connection: Yearbook of the National Society for the Study of Education* (pp. 131-152). Chicago: University of Chicago Press.

Sperling, M., & Freedman, S. W. (2001). Research on writing. In V. Richardson (Ed.), *Handbook of Research on Teaching* (4th ed., pp. 370-389). Washington, DC: American Educational Research Association.

Vygotsky, L. S. (1978). *Mind in society: The development of the higher psychological processes*. Cambridge, MA: Harvard University Press.

Witte, J. A., & Faigley, L. (1981). Coherence, cohesion, and writing quality. *College Composition and Communication*, 32, 189-204.

Sherry Seale Swain is a senior research associate in the Research and Evaluation Unit of the National Writing Project, working with small teams of researchers across the country to study writing instruction in K-12 classrooms. Her current research interests include the development of writing assessments, links between assessment and instruction, and the impact of scoring experiences on teachers. Email her at sswain@nwp.org.

Richard L. Graves is professor emeritus of English Education at Auburn University and editor of several volumes on composition and teaching, including three editions of *Rhetoric and Composition*, *Beyond the Cognitive Domain* with Alice Brand, and *Writing, Teaching, Learning*. In addition to prominent feature analysis, his current research interests include the rhetorical and psychological aspects of literacy learning. Email him at griehardl@bellsouth.net.

David Morse is professor of Educational Psychology and professor of Psychology at Mississippi State University, where he teaches statistics, research and psychometrics courses. His research interests include creativity and giftedness, applications of measurement and statistics principles, and effectiveness of educational programs. Email him at dmorse@ra.msstate.edu.

Copyright © 2021 - *The Journal of Writing Assessment* - All Rights Reserved.