To discuss how training programs for new teaching assistants affect them and, more specifically how a training seminar affected the way they grade papers, twelve teaching assistants read about evaluation, participated in six paper grading sessions, and were assigned an experienced teaching assistant who checked sets of marked essays for validity and reliability of comments and grades. To assess the effects of the training program, data were gathered in four ways: (1) teachers graded the same essay at the beginning and at the end of the semester; (2) each wrote a brief analysis comparing the two gradings; (3) three students from each teacher's class completed a questionnaire about their teacher's grading; and (4) the experienced teachers evaluated the new teachers' grading. Results of training showed that the number, purpose, focus, and format of the comments changed positively, while the grades on the paper decreased. The teachers' responses to their own grading revealed a new confidence in their grading skills and the student questionnaires reflected satisfaction with teacher comments and grades. The evaluations by experienced teachers exhibited concern about inflated grades and unmarked grammatical and mechanical errors. Training sessions for new teachers can influence how they mark papers; however, new teachers themselves may require writing instruction. More research is needed to determine the role of grading in current pedagogy. A student essay for grading by a new teacher is included in the appendices, as well as a coding system for graded essays.

(SRT)
LEARNING TO GRADE PAPERS

In the preface to *Training the New Teacher of College Composition*, Bridges, Lopez, and Lunsford write that "preparation programs vary dramatically because few faculty members involved with training actually have written about their methods, discoveries, or theories.... Until teacher trainers begin to communicate with each other, to work out theoretical concepts about what constitutes good teaching, and to share methods for creating effective instructors, teacher training will remain a hit-or-miss process...." I suspect that even with the contribution of books such as this one, training programs will remain hit-or-miss unless another element is added: research to determine the effects of such programs. It is not enough to collect anecdotal articles telling what we do in training programs and say, "This seems to work." We need some solid investigation and evaluation of our programs.

For the past three years, I have conducted a seminar to train new teaching assistants at Louisiana State University to teach freshman composition. The semester-long seminar meets three hours a week, and we divide the time between discussing composition theory and deciding classroom applications. In part, the new teachers learn on the job since their first teaching assignment accompanies the seminar. They have much to learn including how to guide the writing process, how to structure peer evaluation, and how to
conduct conferences. But one task perplexes all beginning teachers: paper grading. Try as they might to postpone it—with comments on multiple drafts, acceptable/not acceptable ratings, or student conferences—eventually teachers must evaluate the written product, assign a grade, and figure a semester average. Students expect such feedback; our department requires it.

The responsibilities of grading papers come early at LSU when, on the second day of class, all composition teachers assign a diagnostic essay to validate student placement. Then throughout the semester, teachers must assign eight essays. With approximately 25 students in a class, new teachers each grade about 225 essays during the semester. They need help!

But what help can I give them? New teachers would only be unnerved to learn the conclusions that C. H. Knoblauch and Lil Brannon draw from a survey of research on teacher commentary on student writing. They write: "The morose conclusion of [a study by Marzano and Arthur] could well summarize a dominant impression gained from all the research cited here: 'different types of teacher comments on student themes have equally small influences on student writing. For all practical purposes, commenting on student essays might just be an exercise in futility.' Either students do not read the comments or they read them and do not attempt to implement suggestions and correct errors.'" Despite the bleak picture, Knoblauch and Brannon argue that "responding supportively to student writing is indeed central to enlightened instruction." They identify two areas that need improvement: (1) poorly designed experimental studies which attempt to attribute differences
in student performance to written comments that are one small part of a teacher's semester-long interaction with students and (2) traditional commentary that is strictly product-centered and evaluative.

Another "state of the art" article, this one by C. W. Griffin on the "Theory of Responding to Student Writing," gives more hope to new teachers looking for meaningful ways to mark papers. He describes an emerging theory built around three components: our orientations, our verbal responses, and our students' reaction to our responses. A study that I conducted with new teachers shows that as they learn to grade papers, these three components shape their methods.

Description of the Study

Last fall I conducted a descriptive research study to learn how my seminar affected the way new teachers grade papers. Twelve of the twenty teaching assistants in my seminar had never taught; they became my subjects. My training program included readings about evaluation (see Appendix A for a copy of my bibliography), six paper grading sessions throughout the semester in which we compared comments and grades on two or three student essays, and a mentoring program which paired new teachers with veteran ones in the department who checked sets of marked essays for validity and reliability of comments and grades. My goal was to train new teachers to do the following:

* to treat a student's writing with respect
* to consider all aspects of a paper, from rhetorical concerns to mechanics
* to apply different strategies when responding to a draft than to a final copy
* to mark errors in an instructive way (no more awks)
To assess the effects of the training program, I gathered data from four sources: (1) Teachers graded the same student essay twice, once at the beginning and once at the end of the semester, to identify changes in their grading procedures; (2) Teachers wrote a brief analysis comparing the two gradings of the same paper; (3) Three students (one above average, one average, and one below average) from each teacher's class were asked to complete a questionnaire about their teacher's grading; and (4) Mentors wrote reports evaluating the new teachers' grading. Let me first explain my findings from these sources; then I will discuss them and draw some conclusions about methods for training new teachers to grade papers.

Grading the Same Paper Twice

The paper that I asked the new teachers to grade, at the beginning of the semester and at the end, was an actual student paper, typed for easy reading. 4 (See Appendix B for a copy of the essay.) I selected this essay because it had both weaknesses and strengths. While there are some glaring grammatical and mechanical errors, serious problems with organization and content are less obvious. The paper also has strengths: the writer cares about her topic, gives numerous examples, varies her sentence patterns, and uses outside sources to support her argument.

What differences did I expect between the first and second gradings of the essay? The readings, paper grading sessions, and mentor program encouraged new teachers to think about why they mark papers, what they focus on, how they make comments, and where they place them. Therefore to compare the two
gradings, I coded comments in four areas: purpose, focus, format, and placement. The various categories in each area are defined in Appendix C. I adapted coding categories from some used in other studies—particularly ones developed by Elaine Lees and Nina Ziv.

To use the categories, I first determined what constitutes a comment. Then two raters coded each comment. In the event of a disagreement, a third rater judged the category. I also compared the number of comments at the beginning and at the end of the semester and the grades assigned.

What changes occurred between the first and second gradings? One clear difference is in the number of comments. At the beginning of the semester, teachers averaged 19 comments per paper; at the end of the semester, they averaged 29. Comments changed not only in number but also in purpose. At the end of the semester, teachers wrote more emotive comments, mostly praise, and made fewer explicit comments or corrections. They chose instead to write implicit comments that prompted students to reexamine their writing.

Focus of the comments changed as well. The attention to grammar, mechanics, spelling, sentence structure, and rhetorical concerns remained roughly the same; however, the teachers' responses to word choice decreased and comments on content increased. Rather than focusing on how the student said things, teachers examined more closely what the student said.

The format of the comments also changed. During the semester, teachers learned symbols to mark errors. Some used the Harbrace numbering system; others developed their own shorthand. By the semester's end, they also
phrased their responses more often as questions. As for the placement of comments, as the number of comments increased, the space for writing them decreased. At the end of the semester, comments filled the margins and sometimes spilled over into the text.

A final difference between the two sets of papers is the grade itself. At the beginning of the semester, the grades ranged from an A- to a C-; at the end, the range was still a six-point spread, but the grades had dropped from B to D+. However, teachers agreed more often at the end of the semester. At the beginning, 1 teacher assigned an A-, 5 scored the paper in the B range and 6 assigned C's. At the end, 4 assigned B's, 7 assigned C's and 1 assigned a D+. Two teachers' grades were higher at the end of the semester, 5 remained the same, and 5 were lower.

**Teacher Analysis of Grading Practices**

Teachers analyzed the differences between their first and second markings. Because they responded in brief essays rather than a questionnaire, it is difficult to categorize and quantify their responses. Nevertheless one theme emerges. Most teachers said that the second time they graded the essay, they attended more to higher order concerns such as content than they did the first time, a finding supported by coding for focus.

What is more revealing in the teachers' responses is evidence of a new self confidence in their grading skills. Comments such as--"The first time my comments were more copious, but less precise. Too many awks!" and "My second grading was more specific about the problems. I gave more suggestions
for revision"—suggest that the teachers have gained confidence in their abilities to respond meaningfully to student papers.

**Student Response to Grading Practices**

I asked three students from each teacher's class to complete a questionnaire on their reactions to teacher comments on graded papers. (84% returned the questionnaire.) I chose one above average, one average, and one below average student from each class, expecting to find a correlation between a student's grade in the class and his assessment of his teacher's grading practices. After all, if you are getting an A, the teacher must be a good grader! I found no correlation. Instead, nearly all (25 out of 30) reported that their teachers graded fairly. Furthermore, most responded that not only did they usually understand their teacher's comments, but they also used them when revising their graded papers or when writing new essays, finding that contradict other studies on the influences of teacher comments.

When asked if they preferred for a teacher make general suggestions about ways to improve sentences or for a teacher to rewrite some to show what he means, the students were divided. About half favored specific examples, with one below average student explaining, "If I knew how to correct the sentence I would have done it the first time." The other half (mostly above average students) said they preferred general suggestions to make them think; "Don't put words in my mouth," one student responded. These findings support the case study conducted by Ziv in which she found that inexperienced writers "responded favorably to explicit cues" in which she gave specific suggestions for revision. Overall, the questionnaires reflect satisfaction with teacher
comments and grades--although more than one student requested that teachers consider "effort" when assigning grades. (Who hasn't heard that plea?)

**Mentor Evaluations of New Teachers' Grading Practices**

Mentors evaluated new teachers at the end of the semester, addressing their report to the assistant director of freshman composition. While no content for the report was stipulated, all mentors chose to discuss grading skills, and all but one mentor reported competent grading by new teachers. Most reports sounded something like this one: "He had a few problems with grading at the beginning of the semester. His grades were inflated on some papers, he was overlooking some major and minor errors, and his remarks included very little positive reinforcement. But as I discussed each of these problems with him, he worked on them and corrected them. By the end of the semester, the only problem I saw was overlooked misspelled words." This example exhibits two concern that several mentors noted: inflated grades and unmarked grammatical and mechanical errors.

**Discussion**

To discuss the findings from my study, I will examine them in relation to the three theoretical components that Griffin identified: teachers' orientations, their responses, and students' reactions to their responses.

In part, a teacher's orientation to paper grading is determined by the roles she adopts: reader, editor, instructor, judge. More comments about content suggest that over the course of the semester new teachers learned to read papers "interactively." Interactive reading, as defined by Ellen Nold,
is a process, in which "the reader makes meaning from the text; he has an active role... And because of his different experiences and expectations and immediate situation, he may perceive a text quite differently than another reader." For example, in one part of the student essay, the writer gives examples of parental put-downs: "You'll never amount to anything!" or "How could you be so stupid?" or "What's with you anyway?" One teacher responded, "Do parents really say these things?" Another commented, "How true!" Readers (and graders) react differently depending on their experiences.

Another change in orientation occurred as teachers switched roles from editor to instructor. At the beginning of the semester, teachers often corrected the student's errors in spelling, mechanics, usage, and word choice. At the end of the semester, the teachers used symbols to signal errors but required the student to correct the text. Comments became instructive rather than corrective.

Eventually, the teacher's role is that of judge when assigning a final grade. But what exactly is being judged? Although teachers made more comments on content at the end of the semester, it is not possible from this study to determine what most influenced their grades. And related research reports contradictory results. One study found that teachers attend most to content; another discovered that form plays a bigger role; and still a third claimed that although teachers think content and organization are most important, mechanical and usage errors affect their grading most. Perhaps what influenced teachers' orientation in this study was experience; grading
over 200 essays throughout the semester gave them a better understanding of freshman-level writing ability.

Clearly a teacher's orientations influence the verbal responses she makes. In this study, teacher response shows three changes from the beginning of the semester to the end, changes in the number, the focus, and the form of comments. How meaningful are these changes? Currently the pedagogical trend is to limit the number of comments to give students a sense of priority and show them where to focus their efforts. The increase from an average of 19 comments to 29 might be a move in the wrong direction. The practice of limited comments, however, may pertain more to evaluating rough drafts than it does to commenting on final papers. As mentioned earlier, research offers limited constructive advice on the kinds or numbers of verbal responses that are most effective for improving student writing. However, in this study emotive comments, praise for a job well done, account for some of the increase in number of comments. Positive comments have been shown to increase students' motivation for writing.

That teachers began to focus on content more often is not surprising, considering the readings they studied. Several articles stressed a hierarchy for examining student writing, beginning with higher order rhetorical concerns and content before attending to mechanics, usage, and spelling. What is surprising is that the number of global comments relating to rhetorical strategies did not increase. Several readings emphasized the importance of clarity of purpose and audience adaptation.

One last change in teacher response, the form of comments, has already been
explained as a shift in orientation from editor to instructor. To some degree, the teachers' increased use of symbols reflects the preference of mentors. Many mentors request that new teachers learn a particular handbook code of marking so that they can spot check grading more easily. However, the best judges of the effectiveness of symbolic markings are the students themselves.

But the students comprise a hung jury. Their questionnaire reports mixed reaction to the symbols. About as many students reported that handbook codes were the most helpful kinds of comments for improving writing as reported such symbols were the least useful—"I never looked them up." Clearly if teachers employ such a system, they must teach their students how to use it and find a way to make them accountable for corrections.

Although teachers' orientations and style of verbal responses lead them to mark papers differently, students interpret longer comments, those that are not mere symbols, as evidence that a thoughtful reader has considered their writing. When asked to describe the tone of the comments their teacher made, students, responded with such adjectives and phrases as, "encouraging," "positive," "concerned," "never made me feel stupid," and "trying to understand my paper."

Regardless of the grade level of the students, most had positive feelings toward their teachers' responses, seeing them as an effort to help them improve their writing, as this student's comment shows: "In high school I always thought college English teachers would be hard and unfair, but I have been proven wrong this year. Thank God!!!"
Conclusions

What conclusions can be drawn from the study? First, training sessions for new teachers can influence how they mark papers. Furthermore, the influences seem to be positive based on graded essays and comments from the teachers, the students, and the mentors. For the most part my goals for the new teachers were realized, although I would like for them to standardize their grading more and respond to rhetorical concerns in student writing. I am now planning ways to achieve these two goals more fully. Training programs must be thought through, with attention to research in composition pedagogy as well as the requirements of a particular English department.

Second, more research is needed to determine the role of grading in our current pedagogy. With an emphasis on teaching the process of writing using multiple drafts, peer evaluation, and teacher conferences, the role that grading plays is ambiguous at best. Much recent literature has directed teachers to be facilitators and coaches, but not much has been said lately about our responsibilities as judges. Furthermore, adopting a pedagogy based on a hierarchy in which higher order concerns receive attention before lower order ones may mean that an instructor never gets around to teaching lower level concerns except through individual instruction on graded essays. We need to understand better how grading can serve both to instruct and to evaluate.

Third, new teachers themselves may require writing instruction, especially in lower order concerns. As mentor reports indicated, new teachers often
cannot spot errors or articulate the sources of a problem. Awkward becomes a
generic phrase meaning, "Something is wrong here, but I don't know what."
Perhaps the seminar should include readings in a handbook as well.

Fourth, we need to learn more about how experienced teachers grade essays.
While this study has shown that inexperienced teacher improve their grading
practices over the course of a semester, it is not clear what kind of a model
they should emulate. Reliability and validity are goals for marking student
papers. But how do experienced teachers achieve them? Perhaps by conducting
protocols of experienced teachers grading the same student papers we might
find some common procedures.

English departments across the country rely on teaching assistants to help
staff their composition courses. Many departments are beginning to meet the
needs of these new teachers through training programs. This study has
attempted to understand the effects of one training program on the grading of
new teachers. Other areas of teacher training require a similar
investigation.
NOTES


4. I doubt that grading the same essay twice gave teachers a "practice effect" the second time since we did not discuss the essay after the first grading. Between gradings, each teacher marked over 200 other papers. Few recalled what grade they had given the paper; some were surprised to learn that they had graded the same paper twice.


7. Griffin summarizes this research in "Theory of Responding..."

8. Knoblauch and Brannon summarize research on positive comment in, "Teacher Commentary on Student Writing," p. 1.
APPENDIX A

BIBLIOGRAPHY ON PAPER EVALUATION


Below is an essay written about mid-semester by a student enrolled in English 1001. Score it in your usual manner, using whatever marks and comments you typically make. Include a letter grade.

Assignment: Describe a specific problem and offer a solution to it. Write your essay with a particular audience in mind.

Words Will Mold a Child’s Self-Confidence

Recently, while shopping at a crowded grocery store, I followed in succession behind a mother pushing her young son in a grocery cart. Being overly helpful, the boy grabbed a box of Pac-Man cereal from the center of the cereal display. The stack of boxes tumbled to the floor and across the aisle full of busy shoppers. The frustrated mother quickly began to pick up the boxes and shouted to her son: "You are the dumbest kid I've ever seen! Why can't you do anything right? Stop grabbing things or I'm going to kill you!" The young boy sunk low into the cart and fought hard to hold back heavy sobs. I wanted to hug him. He hadn’t been spanked, but the emotional abuse was evident on his tear-streamed face.

Everyone can recall numerous situations from their childhood when they felt crushed by verbal assault. Negative labels such as "dummy" and put-downs like "Can't you do anything right" are all too often aimed at children everywhere. Why does our society ignore it? Parents need to be aware of the tremendous impact of the words they choose to say to their children. Self-confidence can be severely damaged, unintentionally, with negative labels and put-downs. No mistake a child can make is ever worth a degrading remark that crushes their spirit and makes recovery unlikely.

The foundation for self-confidence is laid in childhood. Parents have the power to install confidence within their children, or to take it away. An impaired feeling of self-confidence usually never quite heals. Intelligence, beauty and wealth will never bring contentment to children without confidence in their own judgement and abilities. Parents need to honestly listen to the real message they imply with negative labels and put-downs.

Negative labels are the abusive names that degrade children such as "stupid idiot", "spoiled brat", "jerk", "dirty rat", "fool", "weasel", "retard", "nerd", and "slob."

"Mean Dean lived down the street from me when I was a child."
A Glen C. Griffin, M.D., a writer on child care for the Deseret News put it, "Child abuse isn't all physical--words can cripple and kill."
His parents started to call him "Mean Dean" when his younger sister won a beauty contest and got a lot of extra attention. Dean would treat his sister horribly because he was so jealous. Soon the whole neighborhood knew him as "Mean Dean." Dean grew up to become just what his parents told him to be. He was convicted of armed robbery and sentenced to prison as a young man. Dean was a victim of a very bad label.

Parents often describe a typical little boy as a "cute little monster," or call a seven-year-old girl "baby." The parents then wonder why the boy is reckless and aggressive or why the girl is emotionally immature.

Negative labels can be replaced with positive ones. Instead of calling a child "turkey," tell him he is a genius or simply call him by his name.

Put-downs are also common, and they deeply belittle children. A put-down does not refer to a child's behavior; it jabs painfully at his self-worth or confidence. When parents say "Can't you do anything right?", the child may say to himself "no, I guess I can't."

You'll never amount to anything!" "You can't sing!" "Why don't you run faster?" "Where were you when they passed out brains?" "How could you be so stupid?" "You are totally worthless!" "What's with you anyway?" "You're about as graceful as an elephant." "Who do you think you are, anyway?" It doesn't take much of this to convince a child that they are a nobody or stupid or slow or that they will never be anyone special.

Put downs are cruel and self-defeating to everyone, but children are especially vulnerable because they are developing their self-image. When children behave badly, parents should place the criticism on the behavior not on the self-worth. Instead of saying, "You are bad, clean through," say "You have behaved badly."

The words that parents choose actually program children like a computer. When children hear repeated negative labels and put-downs, how will they be programmed? A child will strive to be whatever they get the most recognition for. In his book "Gifts from Eykie," Dr. Wayne W. Dyer says, "Your child's self-image is a direct result of the kinds of reinforcement he or she receives from you."

Words are the most powerful tool parents have to shape their children's self-confidence. Parents need to be aware and replace the negative labels with positive ones, and criticize behavior, not self-worth. When the little boy in the grocery store knocked over the display of cereal boxes, his mother would have been less abusive and more effective by saying, "Johnny, you have knocked over the display, now help me pick it up."

Feelings and emotions are real and natural to everyone. Parents can learn to manage their feelings so they aren't destructive to their children.
CODING SYSTEM FOR GRADED ESSAYS

Liggett

FORMAT OF COMMENT: Identifies how the comment was written:

1. **Symbol:** The comment is a code or an abbreviation indicating an error. Examples: Harbrace numbers (12-A), "sp" for spelling.

2. **Graphic Mark:** The comment is a mark used without further explanation. Examples: circle, arrow, underlining.

3. **Statement:** The comment is a phrase or sentence that explains an error. Examples: "Needs to be supported by an example or fact." "Essay doesn't flow smoothly in certain areas.

4. **Question:** The comment is a question. Examples: "Transitions?" "Was it really just the label or something else?"

5. **Correction:** The comment corrects an error. Examples: Transposed letters to correct a misspelling, deleted words to eliminate wordiness.

PLACEMENT OF COMMENTS: Indicates where the teacher wrote comments. A paper should be coded on a scale from 1 to 3 as defined below:

1. The teacher writes comments almost entirely in the margins, refraining from writing on or over the student's text itself.

2. The teacher writes comments in the margins and occasionally in the text, but marks in the text to do not hinder reading the original text for the most part.

3. The teacher writes comments all over the paper, in the margins and in the text. Marks within the text mar its reading and often change its meaning through strikeouts or added words.

FOCUS OF COMMENT: Identifies the textual level of concern which the comment addresses, as these examples illustrate.

1. **Grammar/usage:** incorrect reference, need adverb form

2. **Mechanics/punctuation:** underline book title, PUNCT.

3. **Spelling:** sp, or misspelled word corrected or circled

4. **Word choice/diction:** "omit" with an arrow pointing to redundant phrase; "Is this the word you really want?" "Was it emotional abuse or verbal abuse?"

5. **Sentence structure:** Is not parallel; use active voice; combine sentences.

6. **Paragraphing/organization:** "Assertion paragraph is effective"; "Transition?" "Why a new paragraph?"
APPENDIX C, cont.

7. Content: "Good point, but not developed. Instead of adding to the essay it is distracts. Either elaborate or cut it out."

8. Global: (Rhetorical advice on audience, topic, logic--addresses the essay as a whole.) "Readers like to be given good advice too, not just railed at." "The passion as evoked by the stories and other examples shines through."

PURPOSE OF COMMENT: Identifies the teacher's intention in making the remark. Although the purpose of all comments is to some degree instructive, distinctions can be made as follows:

1. Emoting: Generally a supportive comment; a pat-on-the-back reinforcement of good writing. A signal that teachers have feelings too as they read a paper. Emotive comments are usually positive, but not always: "Good idea" "Yawn" "Interesting beginning." When coding these, mark the comment +, *[neutral] or -.

2. Correcting: An explicit comment that corrects the student's error or directs the student how to do so (Harbrace number.) Often in the form of a command that suggests there is a "right way." Examples: sp, 12-A, "Quotation marks not really necessary if all are the same label and in a series." (Translation--take out the quotation marks!) "Use a transition here."

3. Directing: An implicit comment that suggests a possible solution without providing answers; there is still a sense that a "right way" to revise the text exists. A directing comment may remind a student of a particular point in an assignment or class discussion. Examples: "What would be a better word here?" "This paragraph is full of main ideas, none of which is fully developed." "Which statement would you say is your thesis? How can you increase your focus on it?" "How can this example be included with a topic sentence to support a main idea?" "I wish you had made use of the sentence combining exercises we did in the text and on the board."

4. Suggesting: A comment, often in the form of a question, that aims at making the student rethink a problem in the text. There is no clearly right or wrong solution. "Was it really the label that made Dean mean, or something much more complicated. Consider the cause/effect statement that you make here."

5. Clarifying: A comment that seeks clarification or interpretation before the reader can make a judgment. Examples: "What is this word?"

6. Unknown: The purpose of the comment is not clear or very ambiguous. Perhaps it is an indecipherable mark or a stray mark. Use this category only as a last resort.