ESSAY EXAMS AND PAPERS: A CONTEXTUAL COMPARISON

RANDALL L. POPKEN

INTRODUCTION

As recent studies by Behrens, Ostler, Eblen, Harris and Hult, and Wallace all show, the essay examination is one of the predominant writing types used in American higher education. However, though descriptive analyses involving academic genres such as articles (Broadhead *et al.*), essays (Anderson), dissertations (Dudley-Evans), reports (Kalmbach), and abstracts (Ratteray) are fairly common, we know very little about features of essay exams used in actual college classes.

Scholars' reasons for not investigating the essay exam are based on at least two assumptions that writing teachers often make. One is that essay exams are actually bad for students; many teachers see that exams force students to write under extreme time pressure and assume, because students can't prewrite and revise extensively, that exams work against students' mastery of writing process skills. A second assumption is that the essay exam is not really a genre anyway but rather is essentially the same thing as (or even a pale imitation of) the essay students learn in freshman composition. A pedagogical extension of this second assumption is that students automatically learn how to write essay exams while mastering other kinds of writing tasks in freshman composition.

To test the accuracy of these two assumptions, I investigated essay exams in actual college courses. I used a composite of

research methods to compare the rhetorical context of essay exams with that of writing tasks I called "papers"—that is, assignments such as research pieces, done out of class. One method involved surveying fifteen Tarleton State University faculty members from nine disciplines about how they evaluate their essay exams and paper assignments. I gave these professors a series of evaluation criteria for exams and papers, which they rated on a five-point scale:

1. not a criterion
2. a minor criterion
3. an average criterion
4. a fairly important criterion
5. an extremely important criterion

My second method involved analyzing about 200 exam questions and 25 paper assignments sheets from twelve disciplines; part of analyzing the tasks also involved visiting informally with the professors who designed the tasks. Third, to look at textual features of the two discourse types, I analyzed two sets of student writing. One set consisted of 21 exams from a junior/senior/graduate-level sociology course (Multi-Cultural Civilization in the United States). The other consisted of a paper and two exams each by three students in a junior/senior-level history course (History of the American South); these students had received the same grades on the exams as they had on their papers.

The obvious difference between the two contexts involves time: naturally, writers didn't have the time for invention and revision on exams that they did on papers. But I also found some other—less obvious and extremely important—differences between the two tasks in areas of (1) pedagogical function; (2) prompts; (3) rhetorical function; (4) mechanics; (5) style; and (6) organization.

PEDAGOGICAL FUNCTION

I found a fundamental difference between essay exams and papers to be their pedagogical function—that is, the reasons instructors used them for their classes. Although on the survey all the instructors rated "mastering course concepts" as an extremely important evaluation criterion for both papers and exams, my analysis of the tasks and my discussions with the instructors showed that the nature of this "mastery" was different for both tasks.

On the whole, the paper tasks fit into the "pragmatic writing" category suggested by Arapoff-Kramer in her survey of writing tasks at the University of Hawaii. According to Kramer, the ultimate goal of such tasks is to get students to synthesize: "the writer must go beyond the process of making a series of logical generalizations . . . to the creative application of these generalizations in an altogether new situation" (164). In most cases, the instructors used paper assignments to get their students to apply course concepts and principles to events, situations, or texts that had not been discussed in class. Some tasks required broadly-defined "original research," while others required students to work with one or more "outside readings"—scholarly books or articles. One paper assignment in an introductory political science class, for instance, required students to explore the outcome of a piece of state or local legislation of their choice; they drew on class materials for concepts but then applied those concepts to material gleaned from interviews, observations of public meetings, and newspaper articles. A paper task in a sophomore criminal justice class required students to select a country that had not been dealt with in class and then to research some basic principles of its legal system, drawing on the American system as their point of comparison or contrast. The historian teaching the History of the American South course was displeased when several of his students failed to do more than merely summarize outside reading tasks for his paper assignment; he had expected them to do more evaluative and speculative application of course concepts. So dependent were most paper assignments on concept application, in fact, that instructors only gave them in situations (for certain classes, at certain times in the semester) when they could be reasonably confident that students had acquired those concepts.

On the other hand, my analysis of the writing tasks showed that instructors normally assigned essay exams more frequently and for a narrower pedagogical function: to see how much information from textbooks, lectures, and class discussions their students knew and could reproduce. Some questions called for what Kramer refers to as "report-writing" (164)—that is, repeating a series of acts. Such questions, which were usually shorter and worth a small percentage of the entire exam grade, are illustrated in sample 1., from a geology exam:

1. Discuss how minerals are classified through utilization of their physical and chemical properties.

Other questions (which were longer and worth a greater percentage) resembled the category Kramer calls "expository-writing"—writers had to use "specific facts to support a general assertion" (164), as in sample 2., from a political science exam:

2. Many Americans agree that the lower classes in the U.S. suffer from poverty, ill health, and poor education. They argue, however, that this situation is primarily the fault of the lower classes themselves. Drawing on class lectures regarding alternative explanations of poverty and income inequality, discuss the merit of this argument.

In short, the essay exams I looked at were nearly all what one historian called "thinly-veiled objective tests." Even in the case of the expository exams, professors usually had concepts or facts already in mind which they expected their students' exams to contain. In the Multi-Cultural Civilization exams I looked at, for instance, part of one question (see sample 7. below) asked students to "discuss the impact which slavery had on [the development of the Black family]." All the answers to which the professor awarded at least an 85% contained the same set of facts about the origin of the African family and the patriarchal nature of African families.

These differences in pedagogical function between exams and papers led to differences in sources of content and perhaps even differences in means of invention. On the exams, the source of content was largely memorized information; prompts usually directed students toward generalizations, which the students formed using memorized supporting facts. By contrast, the paper assignments forced writers to come up with their own generalizations, searching beyond their immediate memories into old notes, previously-read texts, and, in some cases, researched materials. Accordingly, invention may even have been different between the two task types. In many cases, the exam tasks fit a category Hairston describes as "self-limited" writing: "before the writer begins to write, she already knows most of what she is going to write or she can easily retrieve the content from memory" (444). On the other hand, the paper assignments were often more like Hairston's "extended reflective" writing category: "The writer begins with little more than an idea and . . . a 'felt sense' of what he

or she wants to say, nor knowing precisely where the paper is going or how the ideas will develop" (445). In the latter, then, invention is more likely to occur throughout the writing process, including during drafting and revising.

PROMPTS

Related to differences in pedagogical functions between the exams and papers are differences in how instructors prompted the two tasks. Prompts are an important part of most academic writing tasks since they serve as a writer's primary point of departure; as Horowitz puts it, "what one does when one writes depends on what one is given to begin with" (453). Prompts for papers were generally quite broad, offering students very little suggestion for direction; instructors clearly expected students to be able to select a topic, narrow that topic, and find a specific topic orientation by themselves. A fairly representative paper prompt is sample 3., used in an introductory-level sociology class:

> 3. Analyze from a sociological perspective a subject such as the following: living in a dormitory, rooming house, or commune; working in a lumber yard, factory, liquor store, grocery store, jail; participating in a rally, demonstration, or political campaign.

A key phrase in this prompt, of course, is "sociological perspective": once students chose their topics, part of their effort on this assignment would have to go into deciding which course concepts applied. By contrast, exam prompts were usually much more specific, with topic already completed, narrowed and direction fully prescribed; sample 4., an exam prompt for the same sociology class as sample 3., illustrates this difference:

> 4. What are some typical traits of "masculine" and "feminine"? Which are biologically affected/established and which culturally? How? Illustrate.

These prompting differences may even necessitate different inventional strategies between the two task types. Since their prompts were broad, invention on papers would require broad, unfocused thinking, especially at early stages, to find a direction. On exams, though, the thinking immediately would have to be detailed and restricted to fit the narrow demands of the questions.

RHETORICAL FUNCTIONS

My analysis of the tasks also showed there to be several differences between exams and papers in terms of written product. One such difference involved the range and type of major "rhetorical functions" (Selinker *et al.*) required by the two tasks. Most paper assignments engaged students in four major functions: summarizing, reporting, evaluating, and analyzing. However, the paper assignments seldom asked students to perform the function of arguing the positions on controversial issues. A fairly typical paper task in this regard was a criminal justice assignment that asked students to observe a criminal justice agency and then outline in detail its functions, drawing on course concepts. While writers on this paper had to report and analyze, they would not have to argue, unless of course they took it upon themselves to introduce something controversial and carry on a debate.

By contrast, the analysis also showed that exams were more likely to ask students to argue. While some exams were like papers in the range of functions they demanded, others presented course issues to students and asked them to take a position. For instance, as illustrated in sample 5., a political science exam engages students in an argument about the prison system:

> 5. A number of purposes have been attributed to the use of prisons: rehabilitation, deterrence, incarceration ("getting the criminals off the streets"), retribution, revenge, punishment, etc. Which one (or ones) of these should be considered the primary purpose (or purposes) and why? (Give reasons.) Given this purpose (or purposes), what kinds of measures in the prison system would be needed to accomplish this purpose (or purposes)? Explain.

Therefore, in demanding more than summarizing, reporting, evaluating, and/or analyzing, the exams required students to command a somewhat wider range of rhetorical functions than did the papers.

MECHANICS

The area of mechanics—grammar, usage, and spelling—also shows differences between exams and papers. The survey questioned professors about spelling and grammar/usage separately,

but results were similar: All of those surveyed rated spelling as at least a fairly important evaluative criterion for papers, but only 50% rated it so on exams. And, all the professors saw grammar/usage as at least an important evaluative criterion for papers, but only 43% did on exams. My analysis of the three history student texts bore out these findings as well. Using an analytical system borrowed from Holder, I found that the three students had greater density of mechanical errors on their exams than on their papers, as shown in Figure 1.

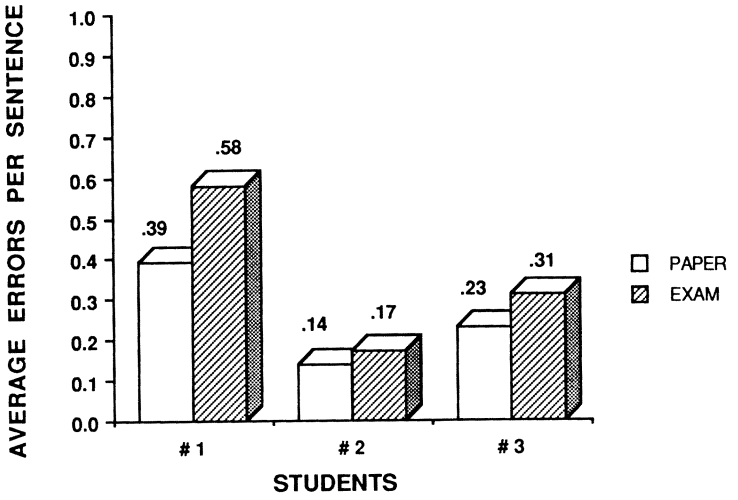

FIGURE 1 - Mechanical Errors on Three History Students' Papers and Exams Receiving the Same Grades

Both of these findings are consistent with informal comments made to me by several professors that they "went easy" on grammar, usage, and spelling problems on exams because they recognized the difficulty many students have editing under time pressure. In short, the threshold for errors appears to be lower on exams than on papers.

STYLE

Differences also appeared to exist between style features that are suitable for papers and exams. One survey question asked

professors whether a "lean style: free of excess words and bureaucratic language" was an important evaluation criterion; 83% of the professors said it was at least a fairly important criterion on papers, but only 54% thought it was so on exams. Again, the professors were making evaluative concessions because of the exam context. However, style differences involving sentence length may have had more to do with writers' responses to the context than the readers' evaluative concessions. Figure 2 shows that the three history students all had shorter sentences on exams than they did on papers.

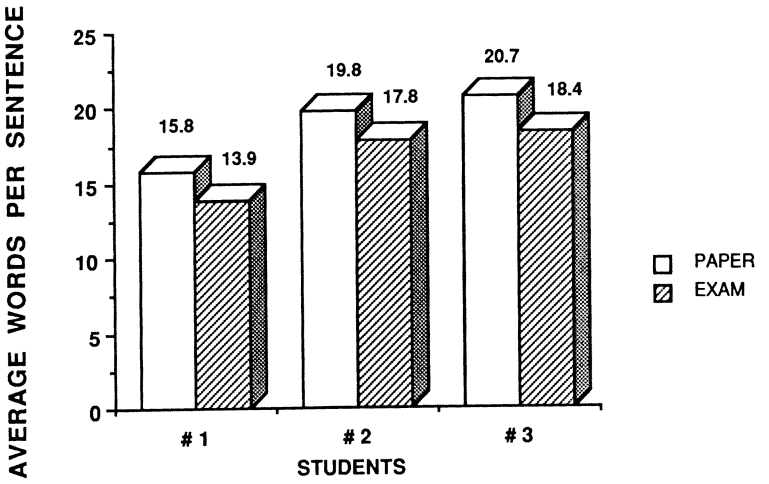

FIGURE 2 - Sentence Length on Three History Students' Papers and Exams Receiving the Same Grades

This is a curious finding since the speed demanded of writing in class doesn't necessarily mean that one's sentences will be shorter (in fact, they could be longer). What is it about the rhetorical context of the essay exam that might result in shorter sentences? On the one hand, the style simply may reflect the writers' psychological responses in composing: as shown in sample 6., the short sentences might be the writer's bursts of short-term memory as she seeks answers to task questions:

6. Congress then passed a series of Reconstruction Acts. Johnson vetoed, Congress overrode the vetoes, thereby taking control of Reconstruction. Congress divided the

South (except Tenn.) into 5 military districts, to be governed by military until a "legitimate" state government was in place.

Or, it may also be that the more skilled writers can even use the shorter sentences strategically. The student in sample 6, perhaps sensing that her professor would be reading rapidly for all the information he could get, may have used the short sentence style (notice too her use of abbreviation) to transmit that information quickly and efficiently, something akin to the style used in a telegraphic message or in a business memo.

ORGANIZATION

Another notable area of difference between the two task types involved organization, at both global and lower levels of text.

In freshman composition and elsewhere, of course, students learn the all-purpose introduction/body/conclusion structure: an introduction sets forth the topic, backgrounds it a bit, and offers a main point that will cover the entire piece; a conclusion provides closure, often reiterating main points in some way. Such introductions and conclusions help give global coherence to a text by making top-level points clear to readers.

However, evidence in two areas suggests that writers use introductions and conclusions differently on exams than on papers. First, on the survey the professors indicated that they had evaluative differences over the importance of introductions and conclusions on the two task types: 91% rated an "introduction section containing a main point (thesis or purpose statement)" as at least a fairly important evaluation criterion for papers, but only 58% did for exams. They regarded conclusions much the same way: 72% rated a conclusion section "that reiterates main points stated in the piece itself" as at least fairly important on papers, but only 41% did for exams. Secondly, the student work I looked at again bore out the survey results. On the history papers, two of the three had traditional introductions, and all three had traditional conclusions; on the exams, none had either. On the Multi-Cultural Civilization exams, only two (of 21) had traditional introductions; none a traditional conclusion.

One reason for so few traditional introductions and conclusions on the exams is related to ways that exams were prompted. It was fairly common for exam questions to consist of several sub-

parts (multiple prompts), sometimes involving three or four different layers. This type of question often does not lend itself to a single macrostatement covering the whole text. The exam prompt from the Multi-Cultural sociology class, as shown in sample 7., was like this:

7. Develop a detailed essay on the Black family. Begin with its origins in Africa and then discuss the impact which slavery had on its development. Continue your essay by discussing the status of the family in the late 19th and early twentieth centuries. Conclude by describing the family in the post-World War II era.

   In your answer, explain the impact which slavery, economics, rural and urban life had on family development. Your answer MUST also include a discussion on Herbert Gutman's study and a discussion on how this study altered thinking about the Black family as developed from previous studies (particularly the theory of E. Franklin Frazier).

   Your answer must also include a major section on contemporary America (as requested in the first section of this question). Discuss the forces which operate today that affect Black family life, emphasizing the growing incidences of single-parent families. In what ways does the Black family reflect or go against the trends of modern America?

Clearly, no thesis or main point is lurking in this question—the response will have to run in several directions, making a global introduction and conclusion very hard to write.

Students' strategies for introducing their answer to question 7. varied. Of the two who attempted global statements, one (the writer of sample 8.) tried to create coherence for her whole text by circumscribing it with the topic "historical & cultural backgrounds of the Black family":

8. In today's news we read and hear about the "epidemic" of single-parent families headed by women and about the skyrocketing rate of teenage pregnancy. These trends both apply especially to Black families. In studying the historical & cultural backgrounds of the Black family unit, we can understand these trends' origins.

The other student resorted to a purpose statement listing topics given in the prompt:

9. In this essay, I'm going to attempt to show the origin of the Black family; the impact slavery, economics, rural and urban life had on family development in the 19th and early 20th centuries. I will conclude my essay by discussing the forces which operate today that affect the Black family.

However, most writers (even ones receiving high grades) disregarded the idea of using an introduction to create global coherence; instead, they took off from one of the sub-points in the question, as the writer of sample 10. did:

10. In Africa, the black family lived in huts. The polygamous male was usually responsible for two or three of these huts within which the female and children lived while he hunted or did other things to sustain the economic support base. The Africans had domesticated crops and animals and a rather orderly way of life which they enjoyed.

Furthermore, the writers of the Multi-Cultural exams used nothing global in their conclusions. Most exams dropped off entirely after the final main point with no real closure. The few that did achieve closure contained phrases such as "no easy solution is in sight to solve the problem" rather than some reiteration of points made in the entire exam.

I also found topic sentences—a conventional means of creating coherence at lower levels to text—to be much less a part of exams than of papers. In the survey, professors indicated that topic sentences were not as important to them on exams as on papers; 100% felt topic sentences were at least a fairly important evaluation criteria on papers, but only 66% did on exams. To verify these perceptions, I submitted the Multi-Cultural exams and the history exams and papers to the minor topic sentence analysis method adapted from Braddock by Popken (211-215). On the Multi-Cultural exams, only 39% of the paragraphs had minor topic sentences; even in the seven exams receiving the highest grades, only 47% of paragraphs had topic sentences, somewhat low by standards in other professional academic writing (Popken, 217). On

the history exams and papers, the differences in topic sentence use were rather remarkable, as shown in Figure 3.:

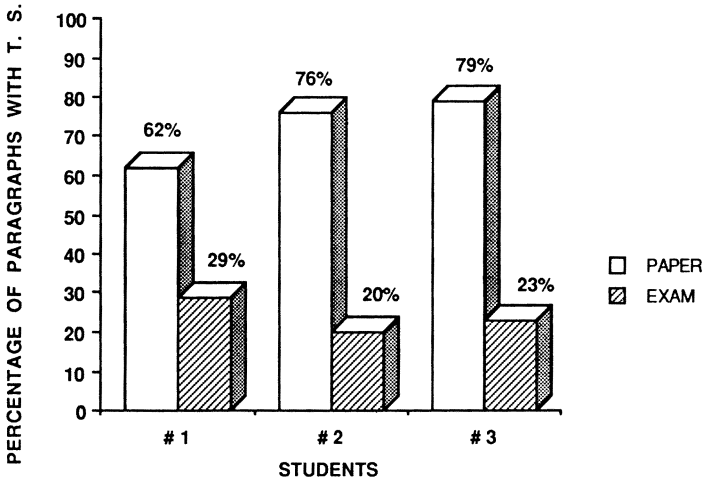

FIGURE 3 - Topic Sentence Frequency on Three History Students' Papers and Exams Receiving The Same Grades

The three history students all relied heavily on topic sentences to stitch their ideas together when they composed papers, but not when they wrote exams.

These organizational differences—both at global and paragraph levels—may, of course, be explained as writers' inability to develop coherence under time pressure. In this sense, the exams might be regarded as examples of the "writer-based prose" that Flower has found in her studies of writers' rough drafts. However, exams are not rough drafts; they are full texts for professor-readers. Thus, the global and lower-level organizational differences have to be seen as deeply embedded in the nature of the two contexts; in short, coherence devices are far less *necessary* on exams than they are on papers. On the exams, the professors brought some very specific presuppositions to their reading—not only regarding information they expected to find but (in the case of the Multi-Cultural professor) even an order for it to appear in. In other words, coherence was already so well established in these readers' minds that introductions, conclusions, and topic sentences were not

generally needed to trigger it; writer and reader, that is, shared something akin to Bernstein's "restricted code" on the exam context. Naturally, when I (or anyone else outside the context) read the exams, they seemed jagged, often without adequate connectives; but the teacher-readers filled in such gaps. On the papers, though, teachers weren't so apt to create coherence for the text; after all, subjects were much broader and content not entirely predictable, or expected. Thus, traditional introductions, conclusions, and topic sentences were more necessary to keep readers abreast of emerging ideas on papers than they were on exams.

CONCLUSION

Evidence from the student writing, the exam and paper prompts, and the faculty survey leads me to challenge the two assumptions about essay examinations noted earlier in this paper. For one thing, exams are not mere "imitations" of papers. Particularly in light of differences in prompts, pedagogical and rhetorical function, style, and organization, exams fit Miller's definition of a genre: "typical rhetorical actions based on recurrent situations" (159). Furthermore, if the exam has generic status, then learning to write essays will not automatically equip students to write essay exams. Maimon has already argued that teaching students only the essay genre is not enough instruction to carry over to reports, summaries, abstracts, research papers and other academic genres they may encounter. Thus, if the essay exam is a separate genre, students also need to be taught how to write it. Moreover, to do so, I believe, means doing more than relegating the essay exam to a "special task" section, near the end of a book, as is common in contemporary rhetorics and handbooks. Instead, in writing courses that attempt to teach academic discourse, the essay exam deserves to share center stage with other prominent academic genres.

But what about the second assumption—the nagging fear that students will somehow develop bad habits from writing essay exams? Admittedly, some students probably conclude from writing exams that they can write everything with haste, but this is more likely to happen if we fail to help them see how their composition strategies need to be flexible, suited to the context demands. Actually, writing teachers are probably as responsible for students' bad habits when they confuse the exam and paper contexts (as

did one English professor who required freshmen to write their final research project—complete with documentation and references—in class). However, the results of this study suggest that the real essay exam as it is used in the disciplines can actually be rhetorically beneficial to students. Essay exams and papers together can give students a fairly comprehensive rhetoric experience and increase their rhetorical flexibility; as the results of this study suggest, the two genres can complement each other in rhetorical demands: While the exams asked students to work with narrowed topics, the papers asked them to work with broader topics and narrow them. While the exams asked students to go head-to-head with specific course issues, the papers asked them to apply course concepts to outside events or texts. While the exams asked students to address an insider audience, the papers asked them to address less knowledgeable readers. Therefore, we can be justified in including essays exams in composition programs—if we can do it in such a way that we teach students to use rhetorical strategies as they approach the exam context. In short, if we teach exams, we should do so by recreating the exam context in our classes, using the prompting, predispositions, and evaluative criteria characteristic of the real genre.

ACKNOWLEDGMENTS

Special thanks to Drs. Michael Pierce and Donald Zelman for letting me work with their students' exams and papers. I am also indebted to Ms. Barbara Schmidt for her work on graphics.

Randall L. Popken is Director of Freshman Composition at Tarleton State University, Stephenville, Texas. His Ph.D. is in rhetoric and composition from the University of Kansas, where he developed his current research interests in text linguistics, writing for academic purposes, and basic writing. His most recent articles appear in the *Journal of Technical Writing and Communication, Written Communication*, and the *Writing Lab Newsletter*.

WORKS CITED

Anderson, Chris. "Essay: Heresay Evidence and Second Class Citizenship." *College English*, 50 (1988): 300-308.

Arapoff-Kramer, Nancy. "A Survey of University Writing Assignments." *College Composition and Communication*, 22 (1971): 116-68.

Behrens, Laurence. "Writing, Reading, and the Rest of the Faculty: A Survey." *English Journal*, 64 (1978): 54-60.

Bernstein, Basil. "Language, Socialization, and Subcultures." *Language and Social Context*. Pier Paolo Giglioli, ed. Baltimore: Penguin, 1972, 163-64.
Braddock, Richard. "The Frequency and Placement of Topic Sentences in Expository Prose." *Research in the Teaching of English*, 8 (1974): 287-302.
Broadhead, Glenn, Marlis Broadhead and James Berlin. "Sentence Structure in Academic Prose and Its Implications for College Writing Teachers." *Research in the Teaching of English*, 16 (1982): 225-40.
Dudley-Evans, Tony. "A Preliminary Investigation of the Writing of Dissertation Titles." *The ESP Classroom: Methodology, Materials, Expectations*. Gregory, James, ed. Exeter: University of Exeter, 40-46.
Eblen, Charlene. "Writing Across-the-Curriculum: A Survey of a University Faculty's Views and Classroom Practices." *Research in the Teaching of English*, 17 (1983): 343-48.
Flower, Linda. "Writer-based Prose: A Cognitive Basis for Problems in Writing." *College English*, 41 (1979): 19-37.
Hairston, Maxine. "Different Products, Different Processes: A Theory About Writing." *College Composition and Communication*, 37 (1986): 442-52.
Harris, Jeannette, and Christine Hult. "Using a Survey of Writing Assignments to Make Informed Curricular Decisions." *WPA: Writing Program Administration*, 8 (1985): 7-14.
Holder, Carol. "Write, Wrong, and the Wred Pen: Understanding Nonstandard Usage in College Composition." *Claremont Reading Conference: 42nd Yearbook*. M. Douglas, ed. Claremont, CA: The Claremont Reading Conference, 202-9.
Horowitz, Daniel. "What Professors Actually Require: Academic Tasks for the ESL Classroom." *TESOL Quarterly*, 20 (1986): 445-60.
Kalmbach, James. "The Laboratory Reports of Engineering Students: A Case Study." *Writing Across the Disciplines: Research into Practice*. Art Young and Toby Fulwiler, eds. Upper Montclair, New Jersey: Boynton-Cook, 1986, 176-183.
Maimon, Elaine. "Maps and Genres: Exploring Connections in the Arts and Sciences." *Composition and Literature: Bridging the Gap*. Winifred Horner, ed. Chicago: University of Chicago Press, 1983, 110-125.
Miller, Carolyn. "Genre as Social Action." *Quarterly Journal of Speech*, 70 (1984): 151-67.
Ostler, Shirley. "A Survey of Academic Needs for Advanced ESL." *TESOL Quarterly*, 14 (1980): 489-502.
Popken, Randall. "A Study of Topic Sentence Use in Academic Writing." *Written Communication*, 4 (1987): 209-228.
Ratteray, Oswald. "Expanding Roles for Summarized Information." *Written Communication*, 2 (1985): 457-72.
Selinker, Larry, Mary Todd-Trimble and Louis Trimble. "Rhetorical Function Shifts in EST Discourse." *TESOL Quarterly*, 12 (1978): 311-20.
Wallace, Ray. "Writing Across the Curriculum: A Faculty Survey." *Writing Lab Newsletter*, 11 (1987): 8-10.