
Listening to Students: A Usability Evaluation of Instructor Commentary

Journal of Business and Technical
Communication

24(2) 206-233

© 2010 SAGE Publications

Reprints and permission:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/1050651909353304

<http://jbt.c.sagepub.com>



Brian Still¹ and Amy Koerber¹

Abstract

Many students see instructor commentary as not constructive but prescriptive directions that must be followed so that their grade, not necessarily their writing, can be improved. Research offering heuristics for improving such commentary is available for guidance, but the methods employed to comment on writing still have not changed significantly, primarily because we lack sufficient understanding of how students use feedback. Usability evaluation is ideally equipped for assessing how students use commentary and how instructors might adapt their comments to make them more usable. This article reports on usability testing of commentary provided to students in an introductory technical writing course.

Keywords

usability, instructor commentary, writing feedback, grading, writing instruction, pedagogy

¹ Texas Tech University

Corresponding Author:

Brian Still, English Department, Texas Tech University, Lubbock, TX 79409.

Email: brian.still@ttu.edu

In her landmark study of instructor commenting practices, Sommers (1982) observed that effective commentary occurs when “what is said in the comments and what is done in the classroom mutually reinforce and enrich each other.” Sommers referred to the mutually reinforcing relationship between the guidance that instructors offer in the classroom and that which they communicate in their commentary on student papers as “the key to successful commenting” (p. 155). Ideally, her perspective suggests, instructors’ comments should provide the same kind of guidance that their classroom lectures and discussion aim to provide—guidance that students can use to improve their writing practices in general, not just to improve their grade on a particular assignment.

Since the time of Sommers’s (1982) study, instructor commenting on students’ writing has received increased attention in our field. Building on Sommers’s early work, scholars have offered a variety of approaches for examining and improving instructor commentary (Brannon & Knoblauch, 1982; Giberson, 2002; Soles, 2001). Some have argued for reflective heuristics meant for instructors to practice or adhere to when providing commentary (Sprinkle, 2004; Straub, 2002). Others have suggested best practices such as recommending that instructors comment only on praise-worthy parts of a paper (Dragga, 1988); forgo editing and simply mark papers as unacceptable, acceptable, or excellent (Dyrud, 2003); or use numbers, instead of letter grades, that can increase with each improved draft (Zigmond, 2006). Although they offer a wide variety of solutions, all of these studies have framed the problem of instructor commenting from essentially the same perspective as that which informed Sommers’s initial study. That is, they have sought to determine how instructors can comment on student writing in ways that they perceive as most beneficial to their students’ long-term success as writers.

In contrast to the extensive research aimed at improving comments from the instructor’s perspective, little attention has been paid to students’ perspectives on commenting. What little research evidence we have, though, has suggested that students see instructor feedback differently from how instructors see it. Whereas instructors see such feedback as constructive criticism aimed at improving student writing in general, students see it, a recent study has suggested, as prescriptive directions that must be interpreted and then followed so that the grade, not necessarily the writing, can be improved the next time (Huot, 2002). Huot observed:

that writing papers for a grade creates a role for the student in which assessing the value of writing is secondary or moot and the attainment of a specific

grade is everything. In this kind of assessment, students are accountable rather than responsible. (p. 168)

Huot's remarks suggest that students' and teachers' expectations for the function of instructor comments just do not match.

To address this persistent problem, some have suggested that writing researchers change their strategies for studying instructor comments. As Cho, Schunn, and Charney (2006) observed, there is "virtually no bibliographic evidence to suggest that instructors, as a general rule, examine, document, or in any systematic way track patterns that emerge in their written commentary" (p. 274). To better evaluate the system, to consider the legitimacy of alternatives, we need not only to research it more; we need, Fife and O'Neill (2001) argued, to focus such research squarely on the "teacher-student exchange" (p. 309). If we take suggestions such as these seriously, we might speculate that the weaknesses that researchers continue to perceive in instructors' comments on student writing indicate the instructors' lack of understanding of their intended audience. But if we perceive this problem as a lack of usability, we might speculate that too much attention has been given to understanding the expert's perspective on this relationship and not enough to understanding that of the user. Some in-depth studies have examined instructor commentary, such as the often-cited study by Connors and Lunsford (1993) that examined comments on more than 3,000 student papers (p. 208), but this research focuses on the nature of the commentary as interpreted through the eyes of experts; such studies have not examined how the intended audience uses and perceives these comments. As Fife and O'Neill observed, "studying only these written texts will not give us all the insight we need." To understand how students use our commentary, we need to engage our students in "conversations and interview[s]"—even though doing so might be more "time-consuming and challenging"—in order to gain a clear, constant awareness of how our audience responds to and uses the critical communication we offer them to become better writers (p. 309).

To address this gap in the previous research, in spring 2007, we conducted a usability study to evaluate the effectiveness of one instructor's comments on a writing assignment from the students' perspective. That is, rather than assessing the comments according to our own preexisting assumptions about what comments should accomplish, we set out to determine how the students were using instructor comments to improve their own writing and how, from their perspective, the comments could be made more useful in this regard. We began by distributing a prestudy survey

distributed to all the students enrolled in four sections of a technical writing service course. As a follow-up to the survey, we implemented a usability evaluation of instructor commentary, focusing on the students. Such an approach has been implemented before for other instructor–student exchanges within the classroom context. Miller-Cochran and Rodrigo (2006), for example, recently assessed the usability of the design of Web-based composition courses. But the usability of instructor commentary on student writing, learned not through self-reporting instruments such as surveys but instead through standard usability testing methods such as close observation and think-aloud protocol, has not been examined before. In this article, we focus on the process by which we conducted this usability evaluation and our initial findings. We argue that although further application of usability-evaluation techniques to instructor commenting might be a fruitful new approach for increasing our understanding of the instructor–student relationship, there are important differences between the teacher–student relationship and the expert–novice relationship that is usually the focus of usability testing. Thus, we recommend how usability-evaluation techniques might be adapted in future research that applies such techniques to instructor commenting.

Specific Study Goals

How do students use instructor comments, and do their actual uses differ from our usual assumptions about students' use of comments? How might the usability of instructor comments be improved through research that seeks to better understand the instructor–student relationship as mediated through instructor comments? These were initially the driving research questions for our study. We later sharpened the study's focus to center on the five key areas that Nielsen (1994, 2003) presented as the defining components of usability:

- *Memorability*. Do students remember the instructor's comments from previous writing, in-class lectures, and examples, and are they able to use them in their current writing?
- *Efficiency*. Are there too many comments to be useful? Can students quickly and effectively use the comments to revise their work without becoming frustrated?
- *Error*. Do students correctly interpret the directions that instructors believe they are communicating?

- *Learnability.* Do future assignments or revisions of assignments indicate that students have learned from the instructor comments?
- *Satisfaction.* Are the comments satisfying? In other words, do the students feel that they have been given what they need to succeed at improving their work?

A great deal of research on usability has been reported in professional and technical communication journals, mostly focusing on user interaction with software, Web sites, instructions, or other types of documentation. But because the context in which instructors, as experts, and students, as novices, relate to each other is not entirely different from the other contexts to which usability-evaluation techniques are often applied, we began with the assumption that usability evaluation is also ideally equipped for assessing how students use instructor commentary and how, given this use, such commentary might be retooled to be made more usable.

Usability testing, especially that which Nielsen (1994) advocated, focuses squarely on contextualized user interaction with the product or process, which Fife and O'Neill (2001) regarded as crucial to understanding how to develop effective commentary. Users are placed in scenarios meant to model those in which they typically operate when using the product or process, and researchers observe them as they perform realistic tasks in those scenarios. While being observed, they are encouraged to think aloud or talk about what they are thinking as they work to complete the tasks. The result, then, is that users do not just tell us what they think; they perform, and we observe, and these different forms of data—from their reported thoughts, their performance, and our observation—when triangulated, provide valuable, in-depth insights into the usability of any number of products and processes.

Other usability-assessment methods, such as heuristic evaluation, can be employed to evaluate a process such as instructor commentary. But heuristic techniques are not, in our opinion, as effective at generating user-driven data as is the approach Nielsen (1994) advocated. Heuristic evaluation, for example, often produces false positive results, meaning that experts conducting the usability assessment find errors that representative users rarely if ever encounter.

We also did not test on location, electing instead to take advantage of a state-of-the-art usability lab equipped with video–audio recording devices, a one-way observation mirror, and other technologies. Arguably, evaluating how users use something in a realistic way is best done in the environment in which they use it, which, for student writers, might be a dorm room or a library. But testing on location often is not feasible, as was the case for our

study. Although the lab where our testing took place was not exactly like a dorm room, it was an environment that we could make consistent for all the study participants. In addition, it allowed us to record video and audio that we could then replay during posttest analysis.

Still, the usability testing we carried out can be done anywhere using much less technology. The key—and what makes it a useful research tool for exploring the effectiveness of instructor commentary on student writing—is to place representative users in representative scenarios and then observe them performing representative tasks. In this case, as explained in the next section, we watched and listened to students as they used the comments they received from their writing instructor to revise a writing assignment.

Methods

We received human-subjects approval for this study from the university's Institutional Review Board (#500739). Our research began with a prestudy survey in spring 2007 (see Appendix A). We surveyed 54 students enrolled in four sections of the same introductory technical writing course. Similar to other survey research on instructor commenting (e.g., Straub, 1997), our research examined how students perceive and use the feedback their instructors provide in response to their writing. Our survey sample included 31 women and 23 men; the average age of these students was 21. When the students, who had taken an average of two college-level writing courses, were asked to self-assess their writing ability using a 10-point scale (10 being the best), the average response was 6.6.

All the course sections that the study participants were enrolled in were taught by the same instructor, who has taught for more than a decade, holds a PhD, and serves as a nontenured lecturer with a 4/4 teaching load at a public university. Typically, each semester the instructor teaches four introductory or advanced professional writing courses. The course syllabus, which is standardized so that the same material is taught for all sections of the course by all instructors assigned to teach it (roughly 10 instructors, 30 courses per semester), covers a range of assignments intended to expose students to the types of writing they would encounter in the workplace, including memos, formal proposals, reports, résumés and cover letters, and instructions. The textbook of choice, Markel's (2003) *Technical Communication*, is supplemented with instructor lectures in class and with online materials (e.g., writing tips, grammar aids, analysis methods, and examples from previous students' work) found on the course Web site and blog.

The assignment that was the focus of this usability study was a formal memo, which was worth 8% of the overall course grade. Students were required to write an informative, analytical, or recommendation memo, 2 to 3 pages in length (or approximately 1,000 to 1,500 words), submitted either in hard copy or electronically.

The instructor painstakingly prepared students to succeed on this assignment as well as others. In addition to offering in-class lectures and supplementary online materials and examples, he required students to complete a memo work sheet before writing the assignment. The work sheet (see Appendix B) forced students to focus their writing intentions before they began writing, thus encouraging the idea that writing is a process. The instructor then responded to these work sheets with suggestions, just as he responded to any drafts that they chose to submit for feedback before submitting their final assignment for grading.

Once the students completed and submitted the assignment, the instructor provided detailed, handwritten commentary on their writing. This commentary (see e.g., Appendix C) represented a mix of in-text notes addressing a variety of microlevel (spelling, grammar) and macrolevel (tone, audience) writing concerns. Further, as a supplement to this in-text commentary, the instructor provided a cover sheet with a grading rubric that listed the seven categories that he considered in determining the overall assignment grade: editing, grammar/clarity, audience, goal, style/readability, formatting/layout, and directions. He placed comments in each rubric category along with a letter grade. The cumulative grade appeared at the top of this sheet as a percentage (see Appendix D for a completed grading rubric).

From the 54 students who participated in the initial survey, we selected a sample of 12 students to participate in user testing of instructor commenting. For this smaller sample, we selected students who were representative of the larger study population. Thus, we chose three students from each of the four sections; six women and six men. This number is more than sufficient (Nielsen, 2000; Nielsen & Mack, 1994) for discovering a significant percentage of the usability problems that exist for a particular user population, such as student writers. All 12 participants were 21 years of age, representing the average age of the participants in the larger study population, and all assessed their writing near the 6.6 average reported in the survey. Finally, the average grade these students earned on the assignment was 77%, which reflected the average grade for all students in the four courses.

The environment in which we asked these 12 students to participate was representative of the one in which they would typically work while revising

Table 1. Severity Scale of Usability Problems

Description
<i>Severely frustrating:</i> The instructor comment cannot be comprehended, so the student cannot use it to understand directions, revisions, or advice that should be taken or considered; such a comment not only takes up time but also ultimately results in failure.
<i>Moderately frustrating:</i> The instructor comment creates significant delay or frustration.
<i>Frustrating:</i> The instructor comment takes moderate effort to comprehend and can be regarded as irritating by the student.
<i>Negligible:</i> The instructor comment presents a challenge of interpretation because of the wording or the handwriting or the location on the paper, but the impact is slight.

their papers. We provided these students with access to computers with Microsoft Word installed on them and told the students that they had 2 hours to use the instructor commentary on their previously submitted memo assignment in order to revise and resubmit it for the opportunity to receive an improved grade.

During this testing, we encouraged students to think aloud (Boren & Ramey, 2000), and we recorded audio of these comments. We also recorded video of the students and used screen capture software to record the work they carried out on the computers to revise their papers. At least two evaluators, including trained student assistants or the study's principal investigator, were present to observe and record the students' comments and work.

For documenting our observations, we created an observation log that allowed us to systematically track and then later analyze why an event occurred, describe it, and then rank it according to its severity. The severity scale we developed (see Table 1) is similar to the one Dumas and Redish (1999) created to indicate the severity of usability problems. Each evaluator worked independently to track the usability problems that participants encountered during the think-aloud protocol, to rank each issue on the severity scale, and to characterize why the problem occurred. Once the evaluators had completed their list of problems, rankings, and characterizations, we consulted the video recordings as necessary to resolve any discrepancies between the evaluators and arrived at the results.

At the conclusion of testing, students were interviewed and given a post-test survey. This interview and survey session was also recorded on camera. The survey offered a mix of closed- and open-ended questions. Using the retrospective recall technique, the interviewer followed up on key

Table 2. Responses to Pretest Survey Questions ($N = 54$)

When writing, how important is instructor feedback (either on your previous writing or on a draft or earlier version of your current writing) to the success of your writing?	Most important: 24 responses Important: 29 responses Not very important: 1 response Not at all important: 0 responses
When considering the following things you rely on to write a document in a class, rank (1 for the best, 2 for the next best, etc.) in order their importance to your writing.	Assignment instructions: 79 Instructor lectures in class, notes: 138 Instructor feedback: 161 Student examples: 180 Assigned readings: 231

comments or actions that we observed during testing. Redish (2007) and Barnum (2002), among others, have noted that this technique is a good supplement to think-aloud protocol, which occurs during actual task completion, because users are not as cognitively overwhelmed in the posttest setting when asked to recall their actions.

Results

In the following subsections, we present the results of the pretest survey, the usability testing, and the posttest survey and interviews.

Pretest Survey

When asked about the value of instructor feedback (see Table 2), all but 1 of our 54 survey participants ranked it as most important (24 responses) or important (29 responses). But when students were asked in the pretest survey to rank the forms of instruction they most relied on to complete their writing assignments successfully, instructor feedback fell squarely in the middle (weighted response: 161). In comparison to instructor feedback, students indicated that they relied more on assignment instructions (weighted response: 79) and lecture notes (weighted response: 138) and less on student examples (weighted response: 180) and assigned readings (weighted response: 231). Thus, our pretest survey results echo the findings of previous research suggesting that students prefer forms of writing instruction that

tell them what to do over more open-ended forms of instruction that require them to determine on their own how best to prepare a written assignment (Huot, 2002).

In other words, these students indicated that they do not rely on individualized feedback as much as they rely on less individualized, but perhaps more directive, forms of instruction such as assignment instructions and instructor lectures. The two categories ranking lower than instructor feedback, student examples and assigned readings, require even more active engagement from the students, again supporting the findings of previous research. Thus, we might speculate that even though the students who participated in this study consider instructor feedback to be important, they are more inclined to rely on less individualized forms of instruction such as assignment instructions and instructor lectures.

Usability Testing

Analysis of the think-aloud protocol of the test, in which we observed 12 student participants as they attempted to use instructor comments to revise their writing assignment, revealed a total of 86 usability problems. Of these 86 problems, evaluators characterized 26 (30%) as severely frustrating and 36 (42%) as moderately frustrating. The remaining problems were characterized as either frustrating (12) or negligible (12).

According to the severity scale we were using (see Table 1), comments characterized as severely frustrating were the ones that students were not able to use, regardless of their importance. On these occasions, the instructor had not clearly communicated with the students, and these unclear comments delayed the students, frustrated them, and led them to bypass these suggestions for revision that they perceived as confusing or potentially more difficult to implement, opting instead just to use those comments that were more discernible or that suggested revisions that were easier to make.

In addition to ranking the severity of the 86 usability problems, we developed thematic categories to account for the variety of reasons why these problems occurred (see Table 3). In the following discussion, we briefly describe each of these categories, starting with the category that was found to be the most prevalent and ending with the category that was found to be the least prevalent.

The most prevalent usability problem occurred when the instructor had made comments about the student writer's tone, grammar, or awkwardness, and students perceived these comments as ambiguous or vague. This category accounted for 34 (40%) of the 86 usability problems. One student

Table 3. Thematic Categories of Usability Problems

Category Description	Number (Percentage) of Problems in Category
Category 1: Vague description of problems in tone, grammar, or awkwardness	34 (40%)
Category 2: Uninterpretable circles, lines, or symbols	30 (35%)
Category 3: Illegible handwriting	11 (13%)
Category 4: Diminishing comments as paper progresses	4 (5%)
Category 5: Marked-out words with no explanation	3 (3%)
Category 6: Inappropriate or insulting tone	2 (2%)
Category 7: Ambiguous or vague underlining	2 (2%)

remarked while revising, “Broad statements, like *awk*, don’t give much of an idea of how to fix something, just that it needs fixing.” Visibly upset, she then gave up and moved to what she described as “easier fixes.” Another student, under the same circumstances, commented, “Most of the time, like I’m doing now, I just fix the easy things fast, the things like spelling that I can figure out that will get me the easy points back. I’m going to take a shot at some of the other stuff, but to be honest, I don’t know what he wants.”

One of the most persistent usability problems in this category involved instances in which the instructor described grammatical or structural problems in terminology that students claimed was unfamiliar to them. Examples of unfamiliar terms that students commented on during think-aloud protocol include the following:

- Seven student participants commented that they did not understand what the instructor meant by the notation *awk*. When asked to clarify this issue in the posttest interview, all of these students claimed that they understood that *awk* was short for *awkward* but that they had trouble discerning why the instructor had marked particular passages as awkward and did not know how to address this comment.
- Two students commented that they did not know the meaning of the term *verb tense*. When asked about this problem during the posttest interview, both students confirmed that they were not familiar enough with this grammatical concept to make the fix themselves and would prefer that the instructor indicate the correct verb tense instead of just pointing out the error.
- One student commented that he did not know the meaning of *pronoun reference*. He confirmed in the posttest interview that he was

unfamiliar with the term and would prefer that the instructor offer a specific recommendation for correcting the problem.

- One student commented that she did not know what the instructor meant by *tone*. When asked about this comment in the posttest interview, she claimed to have a general understanding of the term but did not know how to improve that aspect of her writing.

The second most prevalent usability problem involved circles, lines, or symbols that students claimed they could not interpret. Specifically, 30 (35%) of the 86 usability problems fell into this category. Seven students commented at least once during think-aloud protocol that the instructor had circled a word or sentence with little or no explanation. Six students commented at least once that they did not understand why the instructor had underlined or drawn a line through a particular word or passage.

Four students remarked during think-aloud protocol about symbols or abbreviations that they could not interpret. To illustrate, a student exclaimed, “What the heck is this = thing? I’ve seen this now twice, and he does it on other stuff I do, but I just don’t get it.” Two students commented on the GR notation often used to mark problems with grammar in texts:

He has a blank circle with an arrow coming out of the bottom of it which I don’t know what it means . . . and gr/# which I don’t know how to fix. Is it just grammar, or is there an error with the number? Maybe he means great job [student laughs]. No, just kidding. I’m pretty sure with all these marks everywhere he didn’t mean great job.

Three students commented at least once during think-aloud protocol that they did not understand the location of the problem that the instructor was pointing out in their writing. Specifically, these three students noted during testing and afterward in the posttest interview that they had trouble connecting the instructor’s lines, circles, or other symbols with the specific areas in the text to which the markings were pointing. Also, in some places, the instructor had attempted to use connecting lines in order to present an idea, but the students could not clearly interpret these lines.

Other major problems occurred less frequently but are worth noting because they confirm the overall problems that these students encountered while using instructor commentary. For example, 11 (13%) of the 86 usability problems related to handwriting that participants characterized as illegible. Another problem only represented 4 (5%) of the 86 recorded problems but was nevertheless quite frustrating to students when it occurred: diminishing comments as the paper progressed. Four students remarked about the diminishing number of

comments as their paper neared an end. One student asked, “Does this mean he thinks everything is okay?” Another student, when faced with a final page without comments, said, “Well, if he’s not trying any more, maybe I’ll stop too.”

Final categories for usability problems, each of which represented less than 5% of the 86 problems observed, included marked-out words with no explanation (3 problems), inappropriate or insulting tone (2 problems), and ambiguous or vague underlining (2 problems).

In regard to time or efficiency, the participants took an average of 49.5 minutes to complete their revision of the memo. The longest amount of time taken by a participant was 77 minutes, and the shortest amount was only 24 minutes. When we asked the students if that was how long they typically took to revise an assignment, all 12 said yes. We actually were concerned that the 2-hour limit we set would cause problems for students, but the results indicate that none of the student participants came close to that mark. Students used the instructor comments to find and correct mistakes. When they felt they had done that, they considered their work finished.

Posttest Survey and Interview

The posttesting phase of the study produced mixed results. Although many of the comments that participants made during the posttest interview confirmed and helped to clarify the usability problems that evaluators observed during the think-aloud protocol, some of their written answers to the posttest survey told a slightly different story. Despite the apparently negative remarks that participants made during the think-aloud protocol about the usability of instructor commentary, the multiple-choice items on the posttest survey revealed their generally positive impressions in this regard (see Table 4). Most notably, as Table 4 reveals, 11 of the 12 students rated the comments on their papers as either very useful or useful. Only 1 of the 12 students said the comments were not very useful, and none selected not useful at all. Most participants also positively evaluated the amount and tone of the instructor’s commentary on their writing.

These positive responses on the usability of the instructor comments would seem to contradict the large number of usability problems identified during the think-aloud protocol phase of the study. But the one survey response that might be construed as compatible with the findings from that earlier phase of the study is to a survey item about the amount of time required to use instructor comments. That is, nine students responded that using the comments to make the revisions was very time intensive or time intensive. Only three responded that the time required was short, and none

Table 4. Responses to Multiple-Choice Items on the Posttest Survey (*N* = 12)

1. The instructor comments were	very useful: 3 responses useful: 8 responses not very useful: 1 response not useful at all: 0 responses
2. The tone of the instructor's comments overall was	very positive: 1 response positive: 7 responses negative: 2 responses very negative: 1 response (1 student neglected to circle a response to this question.)
3. The time required to use the comments in order to make revisions was	very time intensive: 3 responses time intensive: 6 responses short: 3
4. The instructor comments' location (where they were placed) was	very short: 0 responses very usable: 2 responses usable: 9 response not very usable: 1 response not usable at all: 0 responses
5. A majority of the instructor comments were	very useful: 2 responses useful: 9 responses not very useful: 1 response not useful at all: 0 responses
6. The amount of instructor commenting was	very adequate: 1 response adequate: 9 responses not very adequate: 1 response not adequate at all: 1 response
7. The instructor comments will	greatly help in the next assignment: 5 responses help in the next assignment: 6 responses have no impact in the next assignment: 1 hurt in the next assignment: 0 responses
8. The instructor comments overall were	very satisfying: 0 responses satisfying: 11 responses not very satisfying: 1 response not satisfying at all: 0 responses
9. Rank the instructor's comments on a scale of 1 (worst) to 10 (best)	average ranking: 7.38

responded that it was very short. This finding is compatible with the previous literature on instructor commenting, our own pretest survey, and many of the comments that students made during the think-aloud protocol and the posttest interview. All this other evidence suggests that students

want forms of writing instruction that tell them what to do to improve their grades. Perhaps their response to this survey item suggests that the participants want instructor comments to help them achieve this goal as quickly and efficiently as possible. Thus, perhaps some of the usability problems revealed during the think-aloud protocol occurred because the instructor comments were not allowing them to accomplish their task as efficiently as possible.

The students' written responses to the open-ended questions on our posttest survey shed some light on this apparent discrepancy between the predominantly negative results of the usability test and the predominantly positive results of the multiple-choice items on the posttest survey. Table 5 lists the open-ended questions included in the posttest survey and a compilation of the written responses provided by 9 of the 12 students who took the time to answer these questions. As Table 5 indicates, the recurring theme in these written responses is that students want instructor comments to be as specific as possible. For example, one student wrote the following in response to question 6, What is the best way for an instructor to comment on your writing? "The best way for the instructor to comment on writing is to give examples of how to fix the problem areas. [The instructor] always just puts comments like awkward or wordy but does not give any clues on how to fix them." Another made a similar remark in response to question 5, Is there such a thing as too many comments? "Comments show you what the prof is looking for." And the same student reiterated this sentiment in response to question 6: "Tell me what you want to see that will give me an A." When they were asked if it was possible for an instructor to provide too much commentary (question 5), only two students said yes, and the rest said no. Many who responded no emphasized this response with exclamation marks, capital letters, or underlining.

Reflecting the overall positive assessment that emerged from the multiple-choice portion of the posttest survey, responses to the open-ended questions revealed a mostly favorable impression of this instructor's comments. Specifically, all of the nine students who responded to question 2 (How successfully did this instructor do that?) on the open-ended portion of the posttest survey made at least one positive remark about this instructor's commenting style. Five of these students added a suggestion for improvement or noted particular comments that they found problematic, but the overall tone of their evaluation was favorable. In sum, then, these written responses to the open-ended questions on the posttest survey suggest that students were generally satisfied with the quality and

Table 5. Open-Ended Questions on the Posttest Survey

1. What makes instructor comments useful or usable?	<p>If you're able to utilize them without being confused Comments are useful when they clearly show a mistake and are usable when an explanation of why it is a mistake is given. Placement is good. The corrections and comments are useful. It made me realize how wordy I am.</p> <p>It's nice when they specifically say what should be changed, but that would be too intensive for the instructor to do 100% of the time. The instructor's comments are useful in that he clearly states what he believes to be wrong with an area of the paper. The only thing that was really confusing was that he was kind of sloppy when underlining and circling problem areas.</p> <p>Helps to understand why I received the grade that I did. Helped me to revise.</p>
2. How successfully did this instructor do that?	<p>Located in column—easy to read. Specific examples, bad & good suggestions on how you did</p> <p>They must be clear, direct, and legible.</p> <p>Correcting mistakes, explaining what was wrong</p> <p>For the most part, okay.</p> <p>[The instructor] did a good job. I found only 3 unclear comments. Very on the first page. Less on the second page.</p> <p>There were many parts that were very specific.</p> <p>As mentioned above he was somewhat successful at this. He could be more successful if he was neater when circling and underlining problem areas.</p> <p>Very well.</p> <p>Very well except w/ the specific examples.</p> <p>For the most part instructions were clear, direct, & legible. I only had trouble identifying one marking.</p> <p>Usually was good and easy to use</p>

(continued)

Table 5. (continued)

<p>3. What makes instructor comments unusable?</p>	<p>If they're unclear When they are unclear or when the comment could pertain to different areas. ([The instructor] wrote a comment, yet I was unsure which sentence he meant.) His handwriting is horrible. I can tell he is in a hurry grading. Broad statements, like AVWK, don't give much of an idea how to fix something, just that it needs fixing. Instructor comments are usable when they identify grammar & spelling mistakes. I really don't get much use out of comments that they make on how to improve my assignment. If they are unclear. Not clear symbols or when their comments are too general. If it is unclear what the comment is identifying or they are hard to read. When he says things are just unclear. It would be nice to know how? Sometimes he just circled or underlined something and didn't specifically write what was wrong. He used his [student drew a picture of a circle] which I don't know what it means & gr/# which I don't know how to fix. (is it just grammar or is there an error with the number) Besides them being messy they are not always clear as to what they are referring to. Nothing specific The instructor clearly identified what was wrong with specific underlined or circled areas of the paper. Some, but very few were hard to read. [Student draws two symbols] unclear of meaning, too general There was one symbol [student draws picture] that I saw twice that I didn't understand, and sometimes comments were hard to place with the text they described. They all were.</p>
<p>4. What specifically did this instructor do that made the comments unusable?</p>	

5. Is there such a thing as too many comments?
- No
NO!!!
NO
- No! Comments show you what the prof is looking for.
Yes, and I believe that he used way too many comments when he was grading my paper.
Not unless you are being timed or have little time.
NO. Any suggestions should be appreciated in an academic environment.
Yes
No!
6. What is the best way for an instructor to comment on your writing?
- Give specific feedback on what they expect or advise
In the format & style [the instructor] uses but neater handwriting & make all comments the same (all on side of page or all in text) maybe make students double space
Consistent comments, perhaps at the end write a explanation.
Tell me what you want to see that will give me an A.
The best way for the instructor to comment on writing is to give examples of how to fix the problem areas. [The instructor] always just puts comments like awkward or wordy but does not give any clues on how to fix them.
Edit and write notes then have session to describe why.
Red markings in the column next to specific examples.
[The instructor's] style is preferable. I don't think that I have seen other comments that look much different, so I am accustomed to this style and understand it well.
-

quantity of comments that this instructor provided. But when prompted to provide specific suggestions for how the instructor could improve his comments, they responded that they wanted him to suggest solutions to the problems he was identifying in their writing rather than just point out the problems.

Conclusions and Implications for Future Research

Echoing the findings of previous research, this study found that students did not seem interested in using the instructor's comments to help them become better writers; rather, they wanted such comments to help them improve their grade on this particular assignment, and they expected that the comments would enable them to achieve that task as quickly as possible. Although this finding in itself might not be surprising to most readers, we argue that the key value of this study is that it reveals without any illusion—without the suspicion that self-reporting data often bring—that students do not use the comments in the way that we think they do. Whereas we might think that students are ignoring our comments, in this study, students who failed to respond to commentary were often confused about the comments' legibility or meaning. As disappointing as these results might seem, they provide an opportunity for teaching, for generating best practices based on data gathered from a methodologically sound usability evaluation. Some of the guidelines detailed here have been forwarded before, but now that they are tied to observed student use rather than just guesswork, their potential for positive impact should be regarded more favorably.

However limited this study may be, it has provided a perspective not previously offered, one that examines instructor feedback in context, from a student point of view. This study's instructor was experienced, and he implemented methods of feedback not different from that employed by most mainstream writing teachers. We could even argue that he went out of his way to aid students more than some instructors do by providing detailed in-text comments, a grading rubric, and revision and regrading opportunities. As clear as this instructor was in the classroom, however, and as helpful as he was in providing explanations, examples, and opportunities for students to submit work to him for feedback, students still perceived many of his comments as unclear, unhelpful, or unusable.

To summarize, this study suggests the following ways in which instructors might be able to improve the usability of their comments from a student's perspective:

1. Avoid using terminology that is unfamiliar to students. In this study, students claimed to be unfamiliar with basic terms that instructors often use to point out problems concerning tone, grammar, and awkwardness. For these students, such terminology was frustrating, and in most instances, they did not attempt to interpret or use such commenting.
2. Make sure that handwritten comments are legible.
3. Steer clear of indistinct or ambiguous circles or lines, and avoid simply marking out or underlining words or sentences with no explanation. Consider providing students with a key that explains your markings or grading an example paper in class, showing students how you evaluate and what the writer could have done to write more effectively.
4. Be sure that the amount of feedback does not diminish as the paper progresses. Students might perceive this lack of commenting later in the paper not as a positive but rather as an indication that you have given up on the document and the student. Consider marking pages that do not need commenting with some sort of acknowledgment (e.g., “good”) that you have read them.
5. Pay attention to placement of comments. Some students reported that they were distracted by comments that were not consistently placed.
6. Offer solutions rather than just point out writing problems.

Some of the suggestions in this list are the ones that all instructors should address. It is hard to disagree, for instance, that we should avoid using terms that are unfamiliar to students, writing illegibly, or using symbols that we have not defined. And we can easily address some of the suggestions by using less specialized terminology, providing a legend that defines symbols, or spending time in class teaching the grammatical terminology with which we expect students to be familiar. A more significant suggestion emerging from this study is that we should recommend solutions to the problems we identify rather than just point out the problems. This suggestion could also be addressed, at least in many cases, without too much difficulty.

Other items on the list, though, might represent areas in which students need to adjust their expectations (and perhaps we need to help them do so by providing more instruction on how to use our comments to revise their writing). The responses offered to the posttest survey item about time might suggest one of these areas. In the multiple-choice portion of the survey, most students expressed dissatisfaction with the amount of time it took to

complete the revisions even though none of the students came close to using the full 2 hours that we allotted for this task. In retrospect, perhaps it was not appropriate to evaluate time as a factor in this kind of study even though time is always a factor in the usability testing of software products. It is doubtful that any writing instructor would want to communicate to their students that revising a paper, no matter how effective the instructor comments, should be done quickly. Addressing this mind-set—that revising is a task that should be done as quickly as possible—is crucial, and focusing more attention on how we comment and how such comments are perceived and used may help us to do so.

This question of time forces us to acknowledge an important difference between the usability testing of a product such as software documentation and that of the pedagogical mechanism such as instructor comments on student writing: Students in the writing classroom are not customers or consumers who are deciding whether to purchase a product. Rather, they are enrolled in a class that is supposed to help them improve their writing. Thus, to assume that instructors should cater to all the demands that students expressed in this study would be overly simplistic. Usability testing of instructor commentary should take this difference into account, especially when considering the time factor. As writing instructors and researchers, we might need to adapt our usual methods of usability testing to the situation of instructor commenting. Although we included the item about time in the posttest survey because the amount of time required to complete a task is a factor that is typically considered in the usability testing of software documentation, in retrospect, this might be one area in which usability testing should be adapted to be most effectively used in pedagogical settings.

The narrow focus of the study might suggest that the findings are too idiosyncratic. In truth, the study was limited to just one instructor, one type of class, and one type of assignment. Future research evaluating a variety of writing courses and instructors, as well as assignments and student populations, is necessary, and we hope our study has encouraged others not just to employ the guidelines this study suggests for more usable commentary but also to pursue further usability testing of instructor commentary. An especially fruitful application for usability testing would be to test alternative strategies for commentary. For instance, the increased use of embedded electronic commenting (Yohon & Zimmerman, 2004), via word processing software such as Microsoft Word, has changed how comments are delivered, but further study is needed to explore whether this new technology for commenting has had a positive effect on the nature of the content. Is it something more usable, or is it the same kind of commenting but just

digitized in call-out bubbles in the margins? Along similar lines, previous work has examined the effectiveness of recording audio feedback (Hunt, 1989; Kates, 1998; Klammer, 1973). Such commentary can be regarded as more conversational than directive, and research has been done (Still, 2006) on its effectiveness when it is embedded into assignments that students have electronically submitted. Usability testing could be a highly effective technique for determining the relative effectiveness of various mechanisms for delivering feedback from instructor to students.

More work must be done to fully examine the usability of instructor commenting. Instructors who are equipped with the right knowledge and guidelines for providing feedback may better aid students in adopting correct writing habits, in changing their perceptions. Still, the problem to overcome may be one of perception more than usability. Our study’s findings suggest that even if we were to make our commentary better, to strengthen the bond between our feedback and our classroom teaching, usability problems might still exist if most students regard usable comments as those that tell them how to get an A—how to fix their grade, not how to write better.

Declaration of Conflicting Interests

The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

The authors received no financial support for the research and/or authorship of this article.

Appendix A

Prestudy Survey

Your Age:	Section No:
-----------	-------------

Gender: [] male [] female

Declared majors:	Minor (if applicable):
------------------	------------------------

No. of writing courses taken (including those in which you are currently enrolled):

(continued)

Appendix A. (continued)

Using a scale of 1-10 (with 10 being excellent), please assess your writing ability/skill

- When writing, how important is instructor feedback (either on your previous writing or a draft/earlier version of current writing) to the success of your writing?
 - Most important
 - Important
 - Not very important
 - Don't take it into consideration or use instructor feedback when writing

- When considering the following things you rely on to write a document in a class, rank (1 for the best, 2 for the next best, etc.) in order their importance to your writing:
 - Assignment instructions
 - Assigned readings
 - Student examples
 - Instructor lectures in class, notes
 - Instructor feedback
 - Other: _____

Appendix B

Memo Preparation Work Sheet

Question	Short Answer
----------	--------------

What is the *specific* subject or issue that your memo will deal with?

What is the *one* most important thing you want the memo to accomplish?

What *specific type* (informational, analytical, etc.) of memo will you write to achieve this goal? What *specific organization strategy* (advantages/disadvantages, temporal, etc.—see Ch. 8) will you use?

Who is your *specific* audience? Ideally, you should be able to name at least one actual person who would read your memo.

How will you appeal to your reader(s)? Briefly describe *at least one* important *specific* quality of this audience and *at least one* *specific* way your memo will reflect this quality.

Appendix C

Student Assignment With Commentary

To: [redacted] *title?*
 From: [redacted] *title?*
 RE: Oil Drilling in the Arctic National Wildlife Refuge
 Date: February 14, 2007

This is a recommendation memo to Senator Kay Bailey Hutchison regarding opening the Arctic National Wildlife Refuge (ANWR) for oil exploration. Oil prices have risen substantially in the past three years. The United States dependency on the Middle East for oil is the cause of this spike in oil prices. The increases in oil prices have a direct effect in our nation's gas prices. Every individual in our nation has felt this price increase financially in one way or another. The opening of ANWR for oil drilling purposes would increase our at home oil production. Our dependency on the Middle East for oil would thus be decreased. This would result in lower fuel prices. The following details information regarding the amount of oil production in the United States, potential drilling areas in ANWR, and how oil production in ANWR could be environmentally friendly.

Much of the reason that the United States is experiencing high oil prices is domestic oil production has been on the fall. Domestic crude oil production has fallen from nearly 9 million barrels per day in 1985 to about 6.6 million barrels per day in 1995. Domestic oil production is estimated to decline to fewer than five million barrels per day by 2010 (www.anwr.com). Due to this we have had to find other sources for crude oil. A large quantity of this oil has come from the Middle East. Secretary of the Interior Gale A Norton was quoted saying, "To avoid unduly betting America's future on Middle East politics, we need to keep finding and opening up new sources of oil and gas." Currently fifty-six percent of our oil is imported and forty-four percent of our oil is domestic (www.solocomhouse.com/anwr.htm). Price gouging is a direct result of oil dependencies from other nations. This dependency is a major reason that the United States is experiencing high oil prices. If ANWR were opened for oil exploration and drilling it would greatly help to reduce this dependency.

It is believed that ANWR is the highest petroleum potential onshore area that has not been explored in North America (www.anwr.com). ANWR has the potential to hold billions of barrels of oil. Oil production in ANWR could rival some of the largest oil fields in the Middle East (www.anwr.com). The main area of interest for oil exploration is less than ten percent of the refuge. This area is believed to hold close to 10.4 billion barrels of oil (www.solocomhouse.com/anwr.htm). In short this means that this one small area of land could with ~~stain~~ *stain* all of America's oil needs for an entire year without importing any oil at all. If ANWR were open for drilling it would increase our domestic oil production significantly. The price of oil could drastically be reduced if ANWR were opened for oil exploration and drilling.

I am aware that one reason the United States government is not pushing for oil exploration in ANWR is the potential thought of drilling fluids polluting water in

migratory waterfowl nesting sites. New environmentally friendly drilling fluids have been created to eliminate this problem. These new fluids are now water-soluble and are broken down in a short amount of time (www.fossil.energy.gov). The water used in the oil drilling process is now safe for the environment as well. The water is no longer allowed to be dumped directly onto the ground after it is used. Water used in oil drilling now goes through a reverse osmosis process along with a series of water filtration systems before it is returned into the environment (www.fossil.energy.gov). After these processes have been completed the water is safe for everything from wildlife to humans.

Many opponents of ANWR being used for drilling also argue that the scenic beauty of the land will be taken away if it is opened for drilling. The new technology of today ensures that the number of pumps can be kept to a minimum. Many of the pumps will be located in locations that will never be seen by humans or wildlife. Today's technology enables a single pump to pull from several different wells. This is achieved by turning a drilling bit vertically or horizontally from the original well and tapping in to other wells that are in a one-mile range (www.fossil.energy.gov). The number of visible pumps will thus be kept to an absolute minimum. This technology has resulted in 22,000 fewer pumps annually across our landscape (www.fossil.energy.gov).

The Arctic National Wildlife Refuge should be opened for oil exploration. This area is known to hold deposits of oil the same size as many of the large oil fields in Saudi Arabia (www.sierraclub.org). The government should no longer keep oil exploration out of this area. The United States need for oil continues to grow as our population and technology grows. We must explore new sources of oil here at home. Our dependency on foreign countries for oil will continue to grow if domestic oil production is not increased. This will lead to oil prices continuing to skyrocket.

Work Cited

Predeger, David (2005). The Arctic National Wildlife Refuge. Retrieved February 11, 2006, from <http://www.anwr.org>

Sierra Club (2005). Just the Facts. Retrieved February 11, 2006, from <http://www.sierraclub.org>

U.S. Department of Energy (2006). Oil and Natural Gas Protection Programs. Retrieved February 18, 2006, from www.fossil.energy.gov

The Ozone Hole Inc (2005). To Drill or Not to Drill, That is the Question. Retrieved March 1, 2005, from www.solcomhouse.com/anwr.htm

edit to get 7 awk sentences

edit

gr/production

tense

Benefits to Ms. Hitchison?
 Clear explanation of
 what you want her
 to do?

Appendix D

Completed Grading Rubric

2.92 = ~~73%~~ (75)



Editing (x 3) Are there serious editing mistakes (typos, spelling, etc)?	Spelling in context. See comments in memo. (B)
Grammar/Clarity (x 2) Are there grammar or clarity mistakes? For ex, are the sentences clear and correct?	Multiple gr and clarity problems. See specific comments in memo. (C)
Audience (x 2) Does the memo have a clear audience? For ex, is it designed for a specific group of people?	Yes. (A)
Goal (x 2) Does the memo have a clear goal that matches the audience?	You have a fairly clear goal, but you do not clearly connect it to the audience. What exactly do you want Kay to do about opening ANWR? She is a busy politician, and will be much more likely to act on a specific concrete recommendation. Vague recommendations make memos much less effective.
Style/Readability (x 1) Does the writing style match the audience? Are there awkward or needlessly complicated sentences? Do you mostly use short, simple, active voice sentences?	Your writing is fairly clear, but there are sentences (especially passive voice) that are longer or more awkward than they need to be. For ex, the sentence "Much of the reason that the United States is experiencing high oil process is domestic oil production has been on the fall" could be reworded as "Domestic oil production has fallen. This has helped to increase US oil prices". Using shorter, more clear sentences would help your audience to read the memo more easily and quickly. (B-)
Formatting/Layout (x 1) Is it in an effective memo format? Is the overall design (font style and size, headings, margins, number of pages, etc) easy for the audience to use?	It is in a clear memo format. However, some clear section headings would probably make it more useful for your audience. (B+)
Directions (x 1) Does the assignment meet the directions? Is it 2-4 pages long? Is there a memo worksheet attached?	Yes. Good job on this. (A)

References

- Barnum, C. (2002). *Usability testing and research*. New York: Longman.
- Boren, T. M., & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. *IEEE Transactions on Professional Communication*, 43, 261-278.
- Brannon, L., & Knoblach, C. H. (1982). On students' rights to their own texts: A model of teacher response. *College Composition and Communication*, 33, 157-166.
- Cho, K., Schunn, C., & Charney, D. (2006). Commenting on writing: Typology and perceived helpfulness of comments from novice peer reviewers and subject matter experts. *Written Communication*, 23, 260-294.
- Connors, R., & Lunsford, A. (1993). Teachers' rhetorical comments on student papers. *College Composition and Communication*, 44, 200-223.
- Dragga, S. (1988). The effects of praiseworthy grading on students and teachers. *Journal of Teaching Writing*, 7, 41-50.
- Dumas, J. S., & Redish, J. C. (1999). *A practical guide to usability testing* (2nd ed.). Portland, OR: Intellect.
- Dyrud, M. (2003). Preserving sanity by simplifying grading. *Business Communication Quarterly*, 66, 78-82.
- Fife, J., & O'Neill, P. (2001). Moving beyond the written comment: Narrowing the gap between response practice and research. *College Composition and Communication*, 53, 300-321.
- Giberson, G. (2002). Process intervention: Teacher response and student writing. *Teaching English in the Two-Year College*, 29, 411-417.
- Hunt, A. J. (1989). Taped comments and student writing. *Teaching English in the Two-Year College*, 16, 269-273.
- Huot, B. (2002). Toward a new discourse of assessment for the college writing classroom. *College English*, 65, 163-180.
- Kates, R. (1998). Tape recorders and the commuter student: Bypassing the red pen. *Teaching English in the Two-Year College*, 25, 21-25.
- Klammer, E. (1973). Cassettes in the classroom. *College English*, 35, 179-181.
- Markel, M. (2003). *Technical communication* (7th ed.). Boston: Bedford/St. Martin's.
- Miller-Cochran, S., & Rodrigo, R. L. (2006). Determining effective distance learning designs through usability testing. *Computers and Composition*, 23, 91-107.
- Nielsen, J. (1994). Heuristic evaluation. In J. Nielsen, & R. L. Mack (Eds.), *Usability inspection methods* (pp. 25-62). New York: Wiley.
- Nielsen, J. (2000, March 19). Why you only need to test with 5 users. *Jakob Nielsen's Alertbox*. Retrieved January 14, 2008, from <http://www.useit.com/alertbox/9710a.html>

- Nielsen, J. (2003, August 25). Usability 101: Introduction to usability. *Jakob Nielsen's Alertbox*. Retrieved January 14, 2008, from <http://www.useit.com/alertbox/20030825.html>
- Nielsen, J., & Mack, R. L. (Eds.). (1994). *Usability inspection methods*. New York: Wiley.
- Redish, J. (2007). Expanding usability testing to evaluate complex systems. *Journal of Usability Studies*, 2, 102-111.
- Soles, D. (2001). Grading as teaching strategy. *Teaching English in the Two-Year College*, 29, 121-134.
- Sommers, N. (1982). Responding to student writing. *College Composition and Communication*, 33, 148-156.
- Sprinkle, R. (2004). Written commentary: A systematic, theory-based approach to response. *Teaching English in the Two-Year College*, 31, 273-286.
- Still, B. (2006). Talking to students: Embedded voice commenting as a tool for critiquing student writing. *Journal of Business and Technical Communication*, 20, 460-475.
- Straub, R. (1997). Students' reactions to teacher comments: An exploratory study. *Research in the Teaching of English*, 31, 90-119.
- Straub, R. (2002). Reading and responding to student writing: A heuristic reflective for practice. *Composition Studies*, 30, 15-60.
- Yohon, T., & Zimmerman, D. (2004). Strategies for online critiquing of student assignments. *Journal of Business and Technical Communication*, 18, 220-232.
- Zigmond, R. (2006). The numbers approach to grading papers. *Teaching English in the Two-Year College*, 33, 296-304.

Bios

Brian Still is an assistant professor in technical communication and rhetoric at Texas Tech University. He previously published an article in *JBTC* on the use of embedded voice commentary to evaluate student writing. He can be reached at brian.still@ttu.edu.

Amy Koerber is an associate professor in technical communication and rhetoric at Texas Tech University. Her research interests include qualitative research, rhetorical theory, feminist studies, and health communication. Her previous publications include articles in *JBTC* and *Technical Communication Quarterly*.