The Genre of Teacher Comments from Hard Copy to iPad

Jennifer Grouling

Ball State University

Although scholars have advocated for new technologies for responding to student work, there has been little study of how commenting style varies across types of technologies. Using a combination of artifact analysis and interviews, this study shows how the comments of five writing instructors varied between hard-copy and iPad-collected papers. Comments were coded for focus and mode based on previous work by Straub and Lunsford (1995). The overall focus, mode, and length of comments remained consistent across types of technology. In addition, the genre of the end comment (Smith, 1997) remained consistent and appeared unaffected by technology use. However, participants made more imperative marginal comments using the iPad. Interviews showed a difference in comfort and tactile experiences with the iPad that may account for this difference. Ultimately, the use of different technologies may affect teachers’ emotions and embodied experiences, which may have a more significant effect on mode of comments than the technology itself. Future studies should further examine the connections between the material use of technology, the emotions of the users, and changes in commenting style.

Keywords: response, commenting, technology, iPad
In the 1980s and 1990s, composition studies worked to both describe current practices in teacher commentary and develop guidelines for best practices (Connors & Lunsford, 1993; Daiker, 1989; Freedman, Greenleaf, & Sperling, 1987; Hillocks, 1982; Schwegler, 1991; Smith, 1997; N. Sommers, 1982; Sperling, 1994; Straub & Lunsford, 1995). Work in this vein has continued in the 21st century (Batt, 2005; Ferris, 2014; Fife & O’Neill, 2001) but has not been as prevalent as in past decades. Current research has examined different technologies for responding to student work. Research on audio feedback has continued (Anson, 1997; J. Sommers, 2012) and has evolved into work on using screencasting to respond to student work (Anson, Dannels, Laboy, & Carneiro, 2016; Thompson & Lee, 2012). Despite scholars’ fruitful discussion of the advantages of audio and screencast feedback, it has not become common practice. Ferris (2014, p. 16) found that of 129 teachers surveyed, all reported using written feedback, and 78% still used handwritten response. However, Ferris claimed great benefits to computer-based feedback: “It is legible, it is clearer and less cryptic, and it is permanent and can be saved for future reference or analysis” (p. 21). Similarly, Griffin and Minter (2013, p. 143) called for composition to rethink pedagogy “for the touch screens of cell phones and tablets.” Using an iPad for response offers the potential to provide handwritten comments with a digital device, and thus might prove to be a better entry point for faculty who are used to hard-copy response. Yet little is known about using tablets for responding to student writing.

How does commenting on a project using a tablet vary from marking hard-copy papers, and is this variance significant? Using a mixed-method approach, this study examines how five instructors’ comments varied between responding to hard-copy papers and using Notability on an iPad. Participants commented on five sample papers per technology and were interviewed about their experiences. Drawing on Straub and Lunsford’s (1995) study of types of teacher response, this article uses artifact analysis to show how the focus, mode, and length of comments compared across technologies. In general, comments did not vary in focus, mode, or length; the one exception was that the mode of imperative comments

was more prevalent on the iPad. When paired with interview data, it became clear that this trend occurred only in the response practices of participants who did not like using the iPad. In concert with Caswell's (2014, 2016) work on response and emotion, I conclude that while the technology of the tablet itself may not lead to significantly different styles of comments, faculty’s emotions related to technology use do influence the mode of their commentary. Thus, we should be attentive to faculty comfort with technologies in both the study of and the distribution of technologies that may be used to evaluate student work.

**Literature Review**

Early studies of teacher comments showed a directive style. Sommers (1982 p. 152) noted the tendency for teachers to “rubber-stamp” comments and appropriate student text. Daiker (1989) reviewed studies showing that marking error was more common than providing praise, and Schwegler (1991, p. 222) called the language of commentary “implicitly authoritarian.” However, as composition pedagogy moved away from a current-traditionalist approach toward a focus on rhetoric, response practice changed. Straub and Lunsford (1995) asked 12 experienced composition scholars to comment on 15 essays. They found that the majority of the comments focused on global issues in the form of questions and reflective statements (Straub & Lunsford, 1995). Straub and Lunsford concluded that best response is well developed, text specific, global, and attentive to the rhetorical situation. More recently, Ferris (2014) showed that these suggestions have been taken up by composition teachers and that response has shifted to focus on specific suggestions rather than correctness. It seems then, that a common philosophy of response in composition has emerged.

This philosophy of stressing praise and global comments is reflected in the genre of the end comment. Connors and Lunsford (1993, p. 210) found that the “most common trope” was to begin end comments positively, then shift to criticism. Similarly, after reviewing 313 end comments, Smith (1997) identified the strongest trend as starting with a positive comment. Batt (2005) agreed and also found that the use of passive voice was paired with negative feedback to soften criticism. Together these studies offer a rich description of the genre of the end comment. While this genre

seems steady, none of these studies took into account the type of technology that was used for response. It can be assumed that the feedback was either handwritten or typed, because of both the date of the studies and the fact that these still represent the most common forms of response. In the last 10 years, we have seen an increase in technological options for response—the iPad was released in 2010—and thus new studies need to be conducted to see how new types of technology might affect the end comment.

Recent articles have discussed how audio and screencasting comments vary in style and length. Sommers (2012) maintained that “recorded commentary differs from written commentary in a meaningful way in kind, that is, in the types of comments made” (para. 3). He suggested three genres of temporal comments that are more common in audio feedback: retrospective, synchronous, and anticipatory comments. Another difference between audio/visual and written feedback is the length of comments. Connors and Lunsford’s (1993, p. 211) analysis found that written comments over 250 words were rare, yet Sommers’s (2012, para. 2) audio comments averaged 830 words. Audio feedback allows for longer comments given in a shorter amount of time and lacks the space constraints of the physical page (Sommers, 2012). While further research and larger sample sizes are necessary, this research represents an important trend in the scholarship of response: the interaction of technology with teacher comments.

Tools for response such as screencasting are unusual and therefore widely studied, but other tools from the purple pen on a hard-copy paper to the comment bubble in a Microsoft Word document have become nearly invisible. Since tablets have become increasingly popular, we also need to think about how our response interacts with these devices. Research has examined reading and writing on tablets, but response has not been the focus of these studies. Computer scientists Ozok, Benson, Chakraborty, and Norcio (2008) found that users had similar levels of satisfaction when using a stylus on a tablet as when typing on a PC despite more potential for writing errors. Writing scholars Mangen, Anda, Oxborough, and Bronnick (2015, p. 230) compared handwriting, typing on a computer, and typing on a digital keyboard, noting that each has “affordances which may influence cognitive aspects at different levels.” Teachers must
consider these affordances when using new technologies for response. For example, Sullivan (2013, p. 2) stressed the potential of the iPad for writing instruction because it is “equipped with various assistive features including a voiceover screen reader, support for closed captioning, and an assistive touch screen.”

However, we must also be aware that these affordances differ based on the individual instructor. In a larger study of iPads in the professional writing classroom, Faris and Selber (2013, p. 390) mentioned that teachers found responding to student work using iAnnotate PDF on the iPad frustrating because of their unfamiliarity with it. They stressed the degree of training and experience of the instructors as a key factor when using the iPad. Similarly, Kim’s (2004, p. 328) study of four instructors found that students’ preferences for the medium of feedback varied significantly depending on the individual instructor, noting that “not all teachers may benefit from commenting in a ‘media-rich’ modality.” Teachers’ experiences and emotions interact with responding to student work (Caswell, 2014, 2016), and our studies of technology and response must also account for these embodied factors.

In order to address this gap, the current study compares teacher commentary on hard-copy papers and teacher commentary using Notability on the iPad. A combination of artifact analysis and interviews was used to answer the following research questions:

1. Do teachers’ commenting styles change when responding to hard-copy papers versus papers on an iPad? If so, how do the focus, mode, and length of these comments shift?
2. What are instructors’ experiences using different technologies for response?
3. Do these experiences account for differences in commenting style between hard-copy and iPad papers?

**Methods**

**Background and Participants**

This study examines the way that individual instructors’ commentary shifts between different technologies as well as instructors’ experiences using different technologies for commenting. In order to do this, I asked
five writing program instructors at a midsized state university to comment on 15 sample papers using different forms of technology and interviewed them about the process. Creswell (2002) explained that mixed methods are appropriate when collecting multiple types of data and provide a more complete understanding of the problem. Like Prior’s (1995) mixed-method studies, I felt that textual evidence alone would not allow for a complete understanding of why the instructors’ responses may have shifted. In addition, Phelps (1989) showed there can be a difference between what teachers say about their response practices and what is actually represented in artifacts, thus the interviews alone could also not show how the response style changed. Thus, the collection of comments combined with interviews was important for answering both how the style of comments may have changed and why these differences might exist. I used a sequential mixed-method design (Creswell, 2002) that consisted of an initial screening survey, a protocol where instructors responded to papers, and a follow-up interview relating their experiences with the protocol to their typical classroom commenting practices. The survey provided some quantitative data but was mostly used to screen potential participants for the main study and provide context within the writing program. The coded protocol data provided quantitative data that allowed for counting and comparing the number of focus and mode codes across participants and across technologies. The interviews added qualitative data that helped to explain the quantitative data.

This project started as a collaboration between myself (then a new assistant professor in the program) and a graduate research assistant (then an MA student in rhetoric and composition). Working together, we developed a broad survey about commentary for writing program faculty in order to screen participants for the larger study and gather context about the writing program. Out of approximately 60 instructors and graduate students teaching in the first-year writing program, 22 completed the survey. All used some form of written feedback (as opposed to audio feedback), and 41% collected hard-copy papers. Only two participants had graded using a tablet. However, the writing program mentors and TAs had recently been given iPads as a part of a technology request for the program, making this study timely.
The survey allowed us to use stratified sampling to select participants based on two factors: (a) experience teaching composition and (b) experience using different technologies for responding to student work. Because teacher response styles can shift from dualistic to relativistic to reflective with teaching experience (Anson, 1989), we selected a mix of participants with 1–20 years of teaching experience. We felt that similar stages of learning might apply to experience with new technologies. Thus it was important to get a variety of participants in terms of overall teaching experience as well as in terms of preferred methods of collecting and responding to student work. We also selected participants who favored different means of response in their classrooms so that our data was not skewed by introducing the iPad as new technology to all of the participants. Table 1 gives an overview of the five participants in the study.

Table 1

<table>
<thead>
<tr>
<th>Pseudonym</th>
<th>Position</th>
<th>Degree/Area</th>
<th>Teaching experience</th>
<th>Method of collecting student work</th>
<th>Method of response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Susan</td>
<td>Instructor (full-time NTT)</td>
<td>MA in composition &amp; literature</td>
<td>Over 20 yrs.</td>
<td>Hard-copy</td>
<td>Handwritten</td>
</tr>
<tr>
<td>Emma</td>
<td>Assistant professor (full-time NTT)</td>
<td>PhD in rhetoric &amp; composition</td>
<td>Over 20 yrs.</td>
<td>Hard-copy</td>
<td>Handwritten</td>
</tr>
<tr>
<td>Desmond</td>
<td>PhD student</td>
<td>Rhetoric &amp; composition</td>
<td>2–4 yrs.</td>
<td>Dropbox</td>
<td>Adobe marginal notes, types end comments</td>
</tr>
<tr>
<td>Charlotte</td>
<td>PhD student</td>
<td>Rhetoric &amp; composition</td>
<td>2–4 yrs.</td>
<td>Blackboard</td>
<td>iPad, typed using Pages</td>
</tr>
<tr>
<td>Clark</td>
<td>PhD student</td>
<td>Linguistics</td>
<td>1st year</td>
<td>Hard-copy</td>
<td>Handwritten marginal, Typed end</td>
</tr>
</tbody>
</table>

Funded by the writing program, we offered the five participants $20 to complete a protocol and a follow-up interview. We also acquired an iPad with this funding that could be used by participants in this study who did not already own one.

**Data Collection and Materials**

For the protocol, we composed sample rhetorical analysis papers ourselves. Our assignment was to analyze an advertisement for ethos, logos, and pathos in 500–800 words. We created our own sample papers to control two aspects of the sample. First, all papers would be responding to the same prompt in the same genre, and second, we wrote a mix of high-, medium-, and low-quality papers so that quality could be controlled. Although our own bias affected what we saw as high, medium, or low quality, we incorporated mistakes that we had seen in teaching rhetorical analysis. For example, one of our low papers involved a student simply describing the chosen advertisement rather than analyzing it. In our interviews we revealed that we had written the samples ourselves, and participants were impressed with how well they fit with their own experiences with student-written rhetorical analysis papers. One limitation to this study, however, was the decontextual nature. In an attempt to counter this limitation, Straub and Lunsford (1995) encouraged participants to create rich contexts such as those that students would experience in real classroom contexts. We, too, took this approach in our protocol instructions and some participants did create additional context. For example, Emma noted that she relies heavily on what she already knows about students when commenting on papers, and, despite our removing gender as a factor by labeling the sample papers with only initials, she gendered each of the fake students and addressed them by name in her end comments. In addition, our participants asked the fake students to come to office hours and referred to previous drafts, textbooks, and classroom lessons.

Participants read a combination of five high-, medium-, and low-level papers in hard copy and a different five on the iPad. We selected the iPad for multiple reasons. Locally, the WPA was able to obtain iPads for all new TAs and their mentors. More broadly, iPads have become more available and used in higher education (Graves et al., 2015; Griffin & Minter, 2013;
Sullivan, 2013). We used Dropbox to transfer the artifacts into Notability (an app on the iPad that allows for the annotation of pdfs) on participants’ iPads. Using Notability allowed teachers to use either a stylus or a keypad when commenting. The stylus could be used to make handwritten comments anywhere on the text, including marginal and end comments. The stylus could also be used to make additional markings, such as stars or arrows. The keypad could be used to insert comments that were typed where the user clicked to insert a text box on the document. In addition, Notability provided the ability to highlight text. All of these features allowed some customizability, including changing the color of the text or highlighter. Although Notability has since added a feature that enables audio feedback, as studied by McKeown, Kimball, and Ledford (2015), we were interested only in text-based comments, as they still represent the majority of teacher feedback (Ferris, 2014). However, part of the analysis involved looking at when instructors chose handwriting or typing and how the type of technology impacted those choices.

In interviews, participants were asked about their experiences with the protocol and about their own practices in the classroom. This again was designed to counter the limitation of the protocol that decontextualized the instructors’ comments. Interviewers also asked participants how they perceived their own style of commenting and how they felt it varied across technologies. Ferris (2014) used a similar sequential mixed-methods approach that moved from survey data to interviews to analysis of comments on student artifacts. This allowed Ferris to follow up on initial survey results and faculty reasons for their response practices and use of technologies. Similarly, the interview data in my study added the participants’ perspectives on the affordances of different technologies for response and the reasons why their practices may have shifted between technologies.

**Data Analysis**

To analyze the textual data from the protocol, we adapted the coding scheme used in Straub and Lunsford (1995, p. 159). Straub and Lunsford separated their coding of teacher comments into two categories: focus and mode. Focus is the content of the comment, what it refers to (p. 160).
Mode is the way that the information is presented, the tone of the comment (p. 166). In their coding scheme, they offer seven codes for focus and ten for mode. We felt strongly that we wanted an equal number of focus codes and mode codes in order to maintain the same level of variance among the two categories. Therefore, we simplified the modes. In particular, we lumped all question-posing into one code. We also eliminated the code for evaluations since we had one for negative evaluations and one for praise (which we felt captured positive evaluations). In testing our scheme, however, we noticed that some comments made by instructors were simply neutral. The most common example of this is when a reader summarizes a point made by the writer: “You’ve described what the ad contains.” While this could be seen as evaluative, it is not clear whether it is positive or negative. Rather it simply states what the author did, and rather than read more into such statements, we added a code for “neutral” comments. In terms of focus codes, we combined ideas and development to make room for a separate code for “rhetorical awareness.” Although Straub and Lunsford include “Who is your audience?” with extratextual comments, we felt with the increased attention to rhetorical awareness within composition—it is the first of the Council of Writing Program Administrator’s Outcomes for First-Year Writing—that this question warranted a separate code. See Table 2 for our adapted scheme.

Table 2

<table>
<thead>
<tr>
<th>Focus (content)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code</td>
<td>Example</td>
</tr>
<tr>
<td>1. Ideas/development</td>
<td>Can you tell me more about x?</td>
</tr>
<tr>
<td>2. Rhetorical awareness</td>
<td>Your audience needs more information about the ad.</td>
</tr>
<tr>
<td>3. Global structure</td>
<td>This paragraph might work better at the beginning.</td>
</tr>
<tr>
<td>4. Local structure</td>
<td>This sentence is confusing; try putting the transition first.</td>
</tr>
</tbody>
</table>

5. Word choice/wording  Too informal!
6. Correctness  Fragment!
7. Extratextual  See me in office hours.

Mode (tone)

<table>
<thead>
<tr>
<th>Code</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Imperative/command</td>
<td>Go to the Writing Center!!</td>
</tr>
<tr>
<td>B. Advisory</td>
<td>You should add a paragraph here.</td>
</tr>
<tr>
<td>C. Question posing</td>
<td>Do you think your audience would know what this is?</td>
</tr>
<tr>
<td>D. Negative evaluation</td>
<td>Wrong word here.</td>
</tr>
<tr>
<td>E. Praise</td>
<td>Great idea!</td>
</tr>
<tr>
<td>F. Reader Response/reflective</td>
<td>I remember reading about this in a newspaper recently.</td>
</tr>
<tr>
<td>G. Neutral</td>
<td>Here you have an example of personal experience.</td>
</tr>
</tbody>
</table>

I employed four volunteer coders (all PhD students in rhetoric and composition) and coded the data myself, for a total of five coders. After dividing the work, each artifact was viewed by three coders. We coded the artifacts in person, first discussing the codes by scoring two example artifacts. We divided units of comments by new sentences or by changing direction within a sentence. For example, the end comment “You have some good ideas here, but they’re not well developed” was divided at the comma and treated as two separate units. Each unit was assigned a focus code and a mode code. After our scoring session, I compiled the codes and selected a final code that best represented the consensus of the group. If two coders agreed but the third did not, I selected the code that the two agreed on. In the fairly rare case that all coders disagreed, I made a determination by looking at the data again and deciding which code best fit.

In addition, I counted the number of times each participant commented using type, handwriting, highlighting, or other marks, such as

underlining, circles, or arrows. I also totaled the word count for each end comment and each marginal comment. Table 3 shows an example of how I recorded the types of comments made for one participant.

Table 3

*Types of Comments Made by Charlotte*

<table>
<thead>
<tr>
<th></th>
<th>Hard-copy margin</th>
<th>Hard-copy end</th>
<th>iPad margin</th>
<th>iPad end</th>
</tr>
</thead>
<tbody>
<tr>
<td># Typed</td>
<td>0</td>
<td>0</td>
<td>49</td>
<td>5</td>
</tr>
<tr>
<td># Handwritten</td>
<td>58</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td># Highlight</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td># Other</td>
<td>74</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Some comments and codes were more straightforward to work with than others. For example, we often found it easy to identify 1. Ideas/Development and C. Question-Posing, a common combination. However, marks that were less clear on the papers were, naturally, harder to code. For example, was “Assumption!” a matter of an idea or a specific word choice? And was the comment an imperative or a negative evaluation? Nevertheless, we reached comparable rates of reliability to other studies. Using Straub and Lunsford’s (1995, p. 393) method of determining interrater reliability that gives .5 for the use of similar codes and 1 point for each exact agreement, I calculated that we agreed 75% of the time on focus codes and 80% of the time on mode codes.

We coded 416 units of commentary on the iPad and 434 units from the hard-copy papers for a total of 1,700 codes. I ran a chi-square test to determine if differences between the codes from the hard-copy papers and the iPad papers were significant. This test allowed me to draw conclusions across the data and answer Research Question 1 regarding whether or not there were differences in commenting style between the two technologies. While we collected a large enough sample of comments to be statistically significant for our participants, to extensively define the way the genre of commentary shifts with different technologies would require an immense sample size, which was not possible with the limited scope and funding of

this study. Lunsford and Lunsford (2008) also explained that it is harder to collect the volume of data now than in the 1980s due to the need for IRB approval at multiple universities. However, since the genre of the end comment—particularly the move to begin with praise—has been well established, I totaled the number of end comments for each instructor that began with praise to compare with Smith’s (1997) genre analysis. Since Smith’s artifacts were not collected via the iPad, this allowed me to see if the same end-comment style held up across technologies.

An advantage of a small sample size was the ability to add qualitative interviews, which allowed me to account for individual faculty style and thus eliminate the possibility that certain faculty who are more inclined to use new technology in their classrooms also lean toward particular styles of response, a factor that would be missed by looking at comments alone. In order to answer Research Question 2 about the experiences of the participants when commenting via different technologies, I used descriptive coding (Saldana, 2009) to summarize passages of the interview transcripts. I used this data to describe the experiences of each participant and to draw conclusions about why their response practices may have varied when using the hard-copy papers versus the iPad. After coding each transcript individually, I compared instructors’ experiences using hard-copy and iPad technologies with one another. This allowed me to find themes and common affordances of each commenting method.

In order to answer Research Question 3 regarding the way that participants’ experiences influenced their commenting style, I drew inferences across the data from the coded comments to the descriptions of experience in the interview transcripts. I compared individual commenting data by calculating the total percent of the time each focus and mode code was used for that participant. I then looked to the interviews to see how this fit with the participants’ own perceptions of their response practices and whether or not they themselves had noticed a shift when working with different technologies. In concert with the coded comments, this allowed me to see the way that different technologies influenced the response of these five participants.
Focus, Mode, and Length of Comments on Hard-Copy and iPad Papers

The overall focus and mode of comments that participants in this study made on artifacts was consistent with previous studies. Straub and Lunsford (1995, p. 182) found that the majority of the comments (55%) focused on global issues, 21% were on local issues, and the remainder were extratextual. They found that the majority were questions and reflective statements (59%), which they labeled as the least directive means of commenting. Only 15% were the most controlling forms such as negative evaluations, imperatives, and corrections (p. 182). Ferris (2014) noted that Straub and Lunsford were not defining typical response patterns but were after ideals from their well-known composition scholar participants. She followed up to see if nearly 20 years later those ideals had become common practice. Ferris found that in both interviews and surveys, “instructors said that higher-order concerns (students’ content or ideas) were more important to them in response” (p. 19). This trend held true in the practice of my participants, regardless of type of technology. In the current study, 64% of all comments were global as opposed to 24% local and 12% extratextual. The focus of these comments did not vary significantly between technologies; the valuing of global feedback remained in keeping with best practice regardless of the technology being used to respond to the paper.

So, too, the genre of the end comment appears to have remained virtually unchanged across technologies. Smith (1997) found that the strongest part of this genre was beginning the end comment with praise regardless of the quality of the paper. Praise was the highest mode overall in end comments in the current study, regardless of technology, with 29% on hard-copy and 27% on iPad end comments. In addition, 16 out of 25 hard-copy end comments and 15 out of 25 iPad end comments started with praise. Those that did not begin with direct praise often began with reader-response or reflective feedback. In fact, only five end comments total began with advice or imperatives. After positive comments, end comments most often led to negative evaluations, advice, or imperatives, which also fits Smith's definition of the genre. Overall, the comments made by the participants in this study appear to be typical of comments made by composition teachers across other studies.
In order to isolate the variable of technology in commenting, I calculated the total times that each code was used, and compared these across hard-copy and iPad comments. I then performed a chi-square test of association. I hypothesized that there would be a shift in commenting style between hard-copy papers and iPad; however, the test was not statistically significant \( (p = .496) \). In only one area was the comparison statistically significant. When comparing the overall number of comments coded with the mode for imperative/command, there were more comments than expected of this type on the iPad than in hard-copy based on the adjusted residual of 2.2. Upon comparing the marginal and end comments on the iPad, it became clear that the greatest part of difference seemed to exist in the marginal comments as indicated by a \( z \)-test of proportions. In all other regards, both the focus and the mode of the comments were similar across technologies.

Comment length also remained consistent. The average end-comment length was similar for those completed on hard-copy papers (73 words) and those on iPads (78 words). The length of marginal comments also did not vary significantly with an average length of 6.5 words on the hard copy papers and 7 words on the iPad. However, it is important to note that these two technologies allowed for different means of writing end comments. One of the affordances of the iPad is offering instructors this choice as it is impossible to type on a hard-copy paper or to use handwriting on many laptop or desktop computers. For the hard-copy papers, four participants handwrote end comments, but one participant decided to type and print his. The reverse happened on the iPad: four participants used exclusively typed end comments, while one preferred handwritten end comments. All marginal comments on the hard copy papers were handwritten; however, they were more mixed on the iPad. Two participants decided to type their marginal comments on the iPad, while the remaining three used the stylus to make handwritten marginal marks. Neither the iPad nor the ability to type comments seems to have impacted the length of marginalia. However, there was a significant difference in the overall number of marginal comments between the hard-copy papers and the iPad. The biggest difference was in the additional markings category. On the hard-copy papers these consisted of arrows, circles, stars, underlining, striking-out,
or punctuation without text (such as ?). The app Notability added in the ability to highlight areas of the text, which was used by those who typed comments. Across all five participants these marks were common on the hard-copy papers, with 218 total; however, for those two participants typing marginal comments on Notability, their other markings dropped from a total of 107 such comments on hard-copy papers to 23 (including highlights) on the iPad. In contrast, those using the stylus on the iPad dropped from 111 other comments to 76, a less striking difference. In this case, the difference in number of other comments may be a factor of typing rather than handwriting comments, although reading papers on the iPad may have had a small effect.

Participant Experiences Responding to Hard-Copy vs. iPad Papers

Differences in the background of instructors is important to consider when looking at the experiences of faculty using different technologies for response. Faris and Selber (2013, p. 379) found that “how students and teachers approached the iPad depended, at least in part, on their attachment and access to other activities and devices and on how comfortable they were with exploring new technology.” This section shows how those differences influenced the current study’s participants as well as their overall experiences using the two technologies.

Susan and Emma had both been teaching the longest and both preferred collecting hard-copy papers and commenting on them using handwritten comments. For these two participants, the tactile feel of hard-copy papers made a difference between the two technologies. Emma talked about the sense of closure she felt when she had completed grading on a stack of hard-copy papers. “It’s physical,” she said. “When I have a paper copy in front of me, I can just flow,” said Susan. “It’s right there; it’s touchable.” Both of these participants were more comfortable with hard copy papers—Susan continually referred to them in terms of comfort food: “mac ‘n’ cheese,” “mashed potatoes.” Both felt that handwritten comments add a certain personal touch. Although some faculty who have always collected hard-copy papers may find moving to the digital distasteful, I was hopeful that since the iPad allowed for handwritten comments completed using a stylus, it would be more accessible to these faculty members.

Susan and Emma’s feelings about the iPad varied significantly. Susan’s opinion of the iPad fluctuated throughout the study. When she first attempted to use it she got frustrated enough that she almost dropped out of the study. However, after becoming more familiar with it, she found that there were some qualities she appreciated. The biggest affordance she saw was that she could zoom in on text, an advantage for aging eyes engaged in hours of reading. She also liked that she could write in purple. On hard-copy papers she had moved to using pencil so that she could erase mistakes, but on the iPad she could write colorful text and still change it. Compared with other digital methods of grading papers, Susan concluded that the iPad most mirrors hard-copy papers, and thus it was her favorite. Emma was also able to maintain her love of handwriting by using the stylus to write comments on the iPad. She made only one typed end comment, just to try it out. Despite this advantage, Emma hated the iPad and stated that she would rather type comments on Blackboard than use the iPad to respond to student work. Emma’s biggest complaint was the lack of marginal space to write in on the iPad. Although Notability has since improved its usability in terms of changing the size of text and the ability to place it more precisely, margin space is still at a premium. One can’t just finish a comment on the blank back of a page, for example. As seen in Figure 1, this led to Emma writing directly over the text, and that countered her philosophy of response and her value of student voice. “That’s sort of diminished their text,” she complained, feeling that she was visually signaling to students that she could take over their work.
Figure 1. Image of Emma’s handwritten comments on the iPad.

Like Emma and Susan, Desmond and Charlotte came from similar backgrounds, and I will discuss them together. Both were rhetoric and composition PhD students with teaching and writing center experience in both their master’s and current programs. Both valued global feedback focused on ideas, both disliked handwriting, and both collected papers digitally. For both, the methods used to respond to papers on the iPad were not completely new. Desmond already collected papers in pdf format using Dropbox, just as the papers on the iPad were gathered. However, he regularly responded using a laptop and Adobe Acrobat comment boxes. Similarly, Charlotte regularly gathered papers digitally. Although she used Pages rather than Notability, she was the only participant who had previously responded to papers on an iPad. While Susan and Emma both felt more comfortable with handwritten comments, both Desmond and Charlotte felt more at home when typing. Both also commented on being able to respond more quickly when typing. Charlotte noted, “I have such bad handwriting that doing the typed comments ended up being a lot better because it was taking me twice as long to write something that was legible.” Comfort is a common theme, and Charlotte was not at all comfortable with handwriting. While Desmond also preferred to type,
he expressed physical discomfort in doing so on the iPad. “I don’t have a physical keyboard for my iPad,” he explained, “So, I have to use the touch, and I don’t like it as well. It’s restricting to me.” Again, there is an important element of physicality to the experience of responding that varies between using hard-copy and iPad papers.

Finally, the new TA, Clark, had just completed his first year of a PhD in linguistics at the time of the study and had taught for only one semester, although he had previously graded papers for a mentor’s class. Clark’s own process represented a combination of hard-copy and digital technologies. He collected hard-copy papers but typed end comments in Microsoft Word. He did the same thing for this study, handwriting marginal comments on the hard copies but insisting on typing the end comment. On the iPad he typed both marginal and end comments. Like Desmond and Charlotte, he expressed concerns about handwriting and felt that he could type “10 times faster.” Clark was early enough in his teaching career that he was still borrowing heavily from the practices of his mentor, Emma. It was by chance that we had a mentor and mentee both participate in the study, but this does indicate the potential for future research on how mentoring impacts response practices. For example, Clark also used grading rubrics created by Emma. This may be why he was collecting hard-copy papers in his own classroom when he expressed a preference for reading on the screen. At the end of the protocol, Clark was excited about the iPad, claiming that the process was much smoother and easier than hard copy.

**The Influence of Experiences on Commenting Style**

In some cases the participants’ experiences coincided with and even explained the changes in focus, mode, and length of comments found through the artifact analysis. Participants were generally aware that they made more nontextual (other) marks on hard-copy papers. Desmond stated that he preferred minimal commenting, and in his regular practice he made few, if any, marginal comments. However, being forced to use hard-copy papers led to a certain reclaiming of marginalia for him. After completing the protocol, he said, “I can see the value of the in-text comments, and maybe I should think of a way to bring that back.” Charlotte, too, found that she missed the affordances of hard copy for making
comments other than text. "I forgot that it's really nice to be able to circle things in hard copy," she said. The two had the highest number of other marginal comments on the hard-copy papers: 74 for Charlotte and 65 for Desmond. Desmond maintained similar types of marks on the iPad and continued to use handwriting for marginalia. However, he cut down the number for these other comments significantly from 65 to 28. Charlotte did not use the stylus to circle or otherwise mark the text on the iPad but did use the highlighting feature in 14 cases. Both saw other marks as an affordance of using hard-copy technology for commenting.

The other major change between hard-copy and iPad artifacts was the increase in the use of the imperative mode on the iPad. While none of the participants specifically mentioned this particular change in tone, they did express experiential differences using the iPad that may have impacted this result. Desmond described his commenting style to me as "questioning," yet his use of questions decreased from 24% of his feedback on hard copy to 19% of his feedback on the iPad while imperatives rose from 6% to 14%. Emma, too, followed the trend of using more imperatives for iPad comments. In addition, her percent of local, sentence-level comments increased from only 2% of her marginal comments on the hard-copy papers to 24% on the iPad, and the amount of praise she included decreased by 10%. While neither characterized their changes of mode in this way, both participants had a sense that they were more negative on the iPad. In particular, Desmond explained in his interview that he felt more critical when using technologies he was less comfortable with, which included the iPad.

When Desmond's feeling of being more critical on the iPad is looked at in conjunction with the artifact data, it is clear that it isn't just a feeling, nor is Desmond alone. Caswell (2014, 2016) studied teacher emotions when responding to papers using think-aloud protocols and interviews. She found that teachers' emotional triggers related to what teachers value (Caswell, 2016) and that emotions were primarily in response to the content of the paper (Caswell, 2014). However, the teachers in Caswell's study were grading papers in their own classroom and were likely comfortable with whatever technology they chose to use in the process. In contrast, the current study shows how affective dimensions and frustrations with learning new technologies can influence the types of comments made by

instructors. Because the papers were not by actual students in the participants’ classes, the emotional connection to the material was lessened, allowing for technology to be the main variable. Clark and Charlotte were both positive about their use of the iPad, and in both cases their degree of imperative comments remained nearly the same across technologies. However, Desmond, Susan, and Emma all expressed frustrations with using the iPad, and in all three cases the number of imperative comments increased by 8–10%. It seems clear, then, that we cannot ignore how the affective embodied experience of response influences the types of comments being made.

Affective differences did not impact the length of comments. Emma’s perception was that she did not write as much on the iPad as she had on the hardcopy papers. However, the length of her end comments increased some from an average of 65 words on the hard-copy papers to an average of 73 words on the iPad, and she made an equivalent number of marginal comments on each (35, 37). Similarly, Desmond felt that he wrote shorter end comments on the iPad due to frustration with the technology, but as with Emma the difference was minimal. His average end comment on hard copy was 62 words versus 57 on the iPad. While both participants’ comments were, in fact, similar in length, their experience was that marginal space felt cramped on the iPad. As seen in Figures 2 and 3, whether typed or handwritten, the marginalia on the iPad does indeed appear close to student text. As the length of marginal comments was not significantly different between the two technologies, this difference had more of an effect on the experiences of the participants than on their actual comments.

Figure 2. Example of handwritten marginal feedback on iPad.
I think it is clear that this ad was made for men because it is kind of sexist. It shows diamonds in the shape of a woman, or at least a bikini. There are two round ones that would be the breasts, and a triangle for down there. This is something that only men would like, but they would see the shape right away. The ad says “She would love expensive lingerie. Happy

**Figure 3.** Example of typed marginal feedback on iPad

Technology is connected to the experiences of its users, and when those users are teachers commenting on student papers, we need to know how their emotional interactions with technologies affect their response. Clearly participants were aware of differences in how they marked hard-copy papers and aware that their frustrations with the iPad led to an overall more negative tone in their comments. However, one of the limitations of the current study was the short amount of time allowed for participants to learn the iPad. It would be interesting to see if imperative feedback decreased if participants used the iPad for a longer amount of time and became more comfortable with the technology. Charlotte, who already had used the iPad for commenting, had no difference in the number of imperative comments. Thus, I suggest that it may be the affective dimension of using a technology and not the technology itself that accounts for this difference in commenting style.

**Conclusion**

This study used a combination of artifact analysis and interviews to examine both how and why instructor commentary varies between responding to hard-copy papers and responding to papers on an iPad. The findings suggest that both technologies offer some unique affordances as well as pose some concerns. Response on hard-copy papers generated more nontextual comments, such as circles, question marks, or arrows. Even though handwriting with a stylus makes such marks possible with the iPad, participants viewed the ability to comment in this way as an affordance of hard-copy technology. Such feedback may be used in conjunction with textual comments in a meaningful way; however, it could also prove confusing for students, who, like the coders in this study, may be unable to tell if such marks are corrections or suggestions.
While less familiar to many instructors, the iPad offers some key affordances that must be considered. The ability to change the size of the text, which was mentioned casually by Susan as an advantage for her aging eyes, is essential for visually impaired students as shown by Henderson, Gibson, and Gibb (2013). In addition, the iPad has been shown to support students with disabilities in reading (Chmiliar & Anton, 2015). Although apps such as Notability still need improvement for translating handwritten text for the visually impaired, the iPad does have the possibility for more accessibility than hard-copy paper. The iPad also offers the affordance of portability, as described by multiple participants as well as Franklin and Smith (2015). The increase in imperative comments is a potential cause for concern, as it is not in keeping with best practice as described by Straub and Lunsford (1995) and Ferris (2014). However, this limitation may be overcome as instructors become more familiar with—and thus less frustrated by—the iPad.

These findings also have possible implications for teacher training. Batt (2005) noted that response is not often talked about explicitly in teacher preparation; however, similarities across teacher response shows that teachers shape their own practice on those they have observed. As seen in this study, Clark based his response practices on those of his mentor, Emma. Desmond, too, mentioned that he began his teaching career collecting hard-copy papers because that is how his mentor had done it. However, he later developed his own way of collecting papers (through Dropbox) and his own minimal style of response. New teachers should be exposed to different types of response, including different technologies. For example, a practicum course might have a workshop that asks students to respond to sample papers using different technologies and discuss the advantages of each. In addition, discussions between mentors and new teachers should go beyond why a particular grade was assigned to why a particular method of response was used.

Training should not end with new TAs, however. It may take time for experienced teachers to adjust to using a different technology for responding to student work, and such shifts should be accompanied by training throughout writing programs. Anson (1989) suggested that response practice becomes more reflective as a teacher gains experience.
However, this study may suggest that such a continuum can be disrupted when unfamiliar technologies are used for response, causing teachers to become once again more dualistic and direct in their feedback. It can be speculated (and is supported by the experiences of Charlotte) that as instructors become comfortable with a new technology, their response once again becomes more reflective.

Future response studies should include technology as a factor. Unless current work is focused on a specific technology for response, such as screencasting, it often does not mention the type of technology used for response. For example, Caswell’s (2014, 2016) recent work on emotion and response represents a promising new direction for response studies but does not connect emotion to technology use. Perhaps this indicates an underlying assumption that comments produced on hard-copy papers are no different than those typed on a virtual iPad keyboard. While the end comment does seem to transcend technologies, there are some very real differences in the ease of making marginal comments in different formats, from the ability to draw arrows to the difficulty of clicking on the right spot to add a marginal comment on an iPad. Thus, the type of technology used for response should be accounted for even in response studies that focus on other elements.

By using a small number of participants in a controlled environment, this study was able to isolate the type of technology used when responding to student work as a factor in the mode of commentary. However, similar research conducted in situ is important for extending this study. One factor not considered is the way that institutional context influences the type of technology used for response. While technology decisions are often up to the instructor, outside pressures can influence these decisions. Writing programs and universities may offer or require certain technologies be used in the classroom. For example, shortly after this study, my institution started asking instructors to use Blackboard Outcomes to align their assignments with certain programmatic goals so that Blackboard could compile student samples for assessment. Thus, more instructors may decide to respond to student work in Blackboard to avoid collecting work in multiple formats. Because of system updates in the middle of data collection, this study did not analyze comments made using Blackboard.

as originally planned, but the effect of course management systems on response is an important area for future study. In contrast to Blackboard, Canvas offers audio and video comment options within the learning management system itself.

Instructors may change their response practices depending on the technologies at their disposal. As new technologies become available to instructors, we need to continue to examine how they impact our response practices and train teachers not only to use these technologies in the classroom but also to respond to student work.

Acknowledgements

This project was funded and supported by the Ball State Writing Program. Alysha Schweigert was an essential collaborator in the development and data collection portion of this project. Katherine Greene, Morgan Gross, Carie King, and Sara Strasser helped code the data. James Jones assisted in conducting the chi-square test. Finally, Daniel Lawson and Tim Lockridge provided valuable feedback on early drafts of this paper. These individuals were essential to the completion of this project and their contributions are greatly appreciated.
References


Thompson, R., & Lee, M. J. (2012). Talking with students through screen-casting: Experimentations with video feedback to improve student