CHAPTER 1.

TEACHERS' BELIEFS ABOUT THE LANGUAGE OF PEER REVIEW: SURVEY-BASED EVIDENCE

Ian G. Anson

University of Maryland Baltimore County

Chris M. Anson

North Carolina State University

Kendra L. Andrews

Wake Forest University

Despite concerns about the lack of research on peer review (e.g., Haswell, 2005), studies of the method have recently been accumulating, especially to test assumptions about the effectiveness of peer review in promoting revision and strengthening students' writing abilities. Recent scholarship, for example, includes studies examining what kinds of comments promote revision (e.g., Leijen; Nelson and Schunn); comparing student and teacher ratings of essays (Moxley and Eubanks; Falchikov and Goldfinch; Cho, Shunn, and Charney); considering peer review from the student's perspective (Brammer and Rees); and tracking what students focus on as they read vs. what they point to when asked to comment (Paulson, Alexander, and Armstrong).

This and other research follows in the wake of decades of instructional advocacy for peer review that links the method with improved writing ability and the development of skills for collaboration (Bean; Spear; Elbow and Belanoff). But it is also clear that peer review involves highly complex cognitive, linguistic, and social-psychological processes that are not always easily employed by novice student writers or taught effectively as part of the writing process. Reflected in the challenges of making peer review work well, these complexities may account for the relatively poor uptake of peer review in higher education, as demonstrated in Braine's survey showing that peer review was the least implemented of the recommended teaching practices for supporting writing. The reasons faculty either gravitate toward or shy away from using peer review are not well known, nor whether specific demographic or teaching-related factors affect these dispositions.

To explore these questions, we conducted a non-probability, voluntary survey of nearly 500 professional writing scholars and educators about peer review. While one goal of the survey was to compare the key terms preferred by teachers to those expressed by students in actual peer review assignments (see Anson and Anson), in our contribution to this collection, we leverage the survey data to more closely examine the dynamics of teacher expectations for peer review. Our goals were first to study the effect of several variables on instructors' attitudes toward peer review, such as workload, institution type, grading load, and academic rank. We questioned if any of these factors related to teachers' attitudes toward peer review and the likelihood that they will employ it in their teaching. Second, we wanted to examine the relationship between teachers' attitudes toward peer review and the key terms, or "quasi-threshold concepts" (Anson, Chen, and Anson) that teachers privilege in response to student writing contrasted with the key terms they think students use in peer review, asking if teachers' attitudes toward peer review related to their faith in students' abilities to use appropriate language reflecting important writing-related concepts. In particular, we were interested in whether the language and concepts teachers privilege in response could provide us with guideposts for how we can encourage the development and use of effective practices for peer review, including the teaching of threshold writing concepts (Adler-Kassner and Wardle).

POSSIBLE SOURCES OF ATTITUDES TOWARD PEER REVIEW

Although the practice of peer review has been popular since the late nineteenth century within academic literary societies and writing clubs, it was not until the 1980s that it appeared in the composition classroom (Gere 304). The process movement brought an intense focus on and interest in revision, with peer review serving as a central activity to promote the improvement of drafts and the learning of rhetorical strategies (see Anson, "Process"). With the subsequent social turn came the introduction of collaborative learning (Bruffee) and the "teacherless writing class" (Elbow, *Writing*), both of which supported in-class activities in which students responded to each other as interested readers and co-creators of meaning. Peer review (also known as peer response, peer editing, or peer feedback) has been a staple in composition classrooms since then, but it has not always been employed in the same way. Some instructors use guided peer-review questions, some have students read their papers to each other, some do a round-table review, some match students in pairs, some keep readers anonymous and some do not, some use word clouds to facilitate response (Illich), and some have students provide their responses using digital peer review systems (Breuch).

Although there is no preferred method for implementing peer review in writing courses, certain principles appear to be consistently valued in the pedagogical literature. Many of these, such as emphasizing constructive criticism and comments that can lead to revisions that focus on global over local concerns while providing a friendly tone, are aligned with similar standards in teacher response (Anson, *Writing*; Elbow, "Closing My Eyes"; Knoblauch and Brannon; Sommers, "Revision"; Sommers, *Responding;* Straub, "Responding"; Straub, *Practice*). Yet the generally positive orientation of the pedagogical literature is not reflected in faculty attitudes toward peer review. The reasons for these disparities are not entirely clear. For example, there may be a reciprocal relationship between the care with which faculty prepare students for and orchestrate peer review, the success they see as a result, and their subsequent attitudes. Weak implementation from a lack of exposure to best practices can lead to poor results, which can, in turn, further diminish faith in the method. In a study designed to gauge the effectiveness of peer review, Charlotte Brammer and Mary Rees administered an end-of-semester survey to students and faculty. The results revealed that "most students find peer review 'not very helpful'" despite how commonly it is used in composition classrooms, pointing to preparation as the key variable for successful peer review (see also George; Graff). Without believing that peer review can realize learning goals, instructors may not invest time in orienting students to the method and helping them succeed (Brammer and Rees 81). The seeming lack of effective revision among peers (and their dislike of the method) convinces the instructors that the time could be better spent doing other things.

In "Peer Editing in the 21st Century College Classroom," Lindsey M. Jesnek uses Brammer and Rees' study as evidence that methods of peer review have not responded to its complexities. While conceding that peer review is an accepted practice supported by research on collaborative learning (Roskelly; Bruffee; Howard; Stewart and McClure), Jesnek points to the challenges facing students when they engage in the practice, which pushes them toward "peer editing" rather than "peer response." The disappointments instructors experience with the method may be predictable in the context of the conceptual and social requirements of successful peer review, such as navigating uncomfortable positions of ego, authority, and agency, and the cognitive requirements of knowing writing-related perspectives that translate into the kind of rhetorical, linguistic, and structural language used in effective response. In many ways, the complexities of peer review parallel those of teacher response, about which Nancy Sommers writes that "although commenting on student writing is the most widely used method for responding to student writing, it is the least understood" (*Responding* 148).

Research on teachers' own response practices is also relevant to their potential uptake of peer review (see Li and Barnard; Nicol). Across a variety of studies

with different methodological approaches, findings indicate that instructors view the provision of "appropriate" feedback as important, echoing Sommers' conclusion from a major longitudinal study of undergraduate writing at Harvard that "feedback, more than any other form of instruction, shapes the way a student learns to write" (Sommers "Across"). Yet some teachers' "chicken scratch" style commentary, necessitated by their workloads and/or lack of training, is ineffective. In a sample of 48 instructors at one university, most felt that feedback practices were remarkably lacking in quality because of instructors' heavy grading loads. When instructors provided feedback to students, these time constraints led them to write terse snippets of commentary on final drafts—summative, "end-loaded" comments rather than more formative types of feedback (Bailey and Garner). Jackie Tuck shows that instructors are aware of their inability to provide quality feedback stemming from both institutional and personal pressures. Faculty are motivated to provide good feedback, but as institutions demand high-quality teaching while simultaneously increasing their workloads, they are generally unhappy with the results—to say nothing of the dissatisfaction of their students (see Sommers "Beyond"). These results suggest a gap between practices advocated in the field, which emphasize ongoing, formative interactions between instructors and students about their writing, and those the instructors used, which focused on summative comments on final texts.

In the context of this problem, peer review would seem to offer at least a partial solution (Nicol; Thomson; and Breslin). In the absence of formative responses from teachers (which may double the teacher's workload), peer review is positioned as a viable way to generate thoughtful commentary that can precipitate productive revision and improvement. We might predict, then, that instructors with high grading loads, especially those early in their careers and/or on the tenure track, will be among those likely to view peer feedback favorably, as a way to overcome the problems noted by Tuck, Bailey and Garner, and others. At the same time, the attractiveness of peer review may be entirely mitigated if teachers don't have confidence that their students can provide useful peer response to their peers' writing in progress. That confidence is reflected in the kinds of language students use to provide feedback. We would expect faculty with confidence in peer review to include richer content in their description of the terms and concepts used by students, more closely matching the terms and concepts used by instructors when they respond to student writing. In contrast, we would expect those less optimistic about peer review to predict students' use of terms and concepts that less closely match the terms and concepts used by instructors. Our study sought to examine these expectations and relate them to the demographic factors described earlier.

WHAT TEACHERS VALUE IN PEER REVIEW: A SURVEY

To conduct our study, we developed, advertised, and distributed a survey designed to assess writing teachers' perceptions of contemporary response practices, as well as conventional assumptions about the content of peer and instructor feedback to writing. The survey was administered on two large e-mail listservs: WPA-L (recently replaced by writingstudies-L but populated at the time by over 3,500 writing teachers and writing program administrators), and the listserv of the European Association for the Teaching of Academic Writing (EATAW), populated primarily by teachers who support and/or study writing in European higher education, especially in English. After two weeks of deployment, the overall N of responses collected across these two listservs totaled 475: 410 from WPA-L, and 65 from the EATAW listserv. Given the size of these listservs' membership, our rate of response was somewhere between 10 and 15 percent. This nonprobability survey allowed us a first glimpse at the perceptions of writing teachers across institutions, nations, academic appointments, and workloads.

The first part of the survey was designed to measure basic demographic information of our sample. As there are no accurate estimates of the population demographics of individuals in the writing studies community, we collected this information as an exploratory exercise. Some demographic information, including measures of racial diversity, educational attainment and background, and marital status, were not collected due to the preliminary nature of the study (and the need for brevity because of the likelihood of drop-out among our respondents, who were not compensated in any way for their efforts). We focused our demographic measurements on three main categories: basic personal attributes such as age, gender, and country of residence; and institutional information such as academic appointment and workload.

Our principal measures of interest focused on respondents' perceptions of the quality of both peer and instructor feedback at the college level, using closed-ended, five-point question batteries asking for overall appraisals of quality. These questions asked respondents if they found peer feedback to be "extremely," "fairly," "moderately," "only a little," or "not at all" helpful to students, and if college and university instructors across the disciplines provide "very high-quality," "high-quality," "moderate-quality," "low-quality," or "very low-quality" feedback on written assignments.

The survey also asked respondents to provide open-ended content describing the terms they expected to find in the feedback of teachers and novice student writers. Instructors' terms, we theorized, represent underlying "threshold concepts" important to the development of effective writing (see Adler-Kassner and Wardle). As a result, the terms respondents believe that students use when

providing peer review would then show the "distance" from the teachers' terms and potentially reflect respondents' experiences using peer review and their perceptions of what students typically know and bring to the peer review process. We first asked respondents to think about concepts that are important for high-quality responses to writing, providing them with ten open-ended text boxes. Next, we asked respondents for terms that might be likely to appear in novice students' responses to writing, providing them with ten additional open-ended text boxes. This data-collection strategy allowed teachers to input their expectations about the type of key lexical content encapsulated in writing-related terms that might appear across a variety of writing assignments among both students and teachers. While this strategy affords a greater degree of generalizability in the lexicons leveraged by our respondents when compared to responses to specific writing prompts, it also potentially widens the range of psychological referents used by respondents, rendering our lexicon quite diverse. Our approach nevertheless resulted in several terms appearing regularly across respondents, meaning that we have likely captured a baseline set of concepts that teachers perceive as important regardless of context. Future studies could use this latter approach to examine teacher response lexicons for more targeted comparisons.

WHO RESPONDED?

Before examining differences across respondents' portrayals of feedback, it is important to first consider the demographics of the sample. Table 1.1 presents basic descriptive information that allows us to assess the nature of our sample of teachers.

An important insight to emerge from Table 1.1 is that our sample of teachers, though predominantly based in the United States, is highly diverse in terms of academic position, workload, and demographics. For example, while roughly 72% of the sample is female, ages range from 25 to 77, with a standard deviation of almost 12 years. We have captured a cross-section of the field that includes members of the composition community at many points in the academic lifecycle—7% of the sample identifies as graduate students, while 11% are untenured tenure-track faculty, 20% are tenured associate professors, and 17% are full professors. Other members of the sample identify their academic position as full- or part-time instructors, administrators, or other appointments; no one type of academic position predominates.

In addition, while the majority of respondents (82%) are responsible for teaching composition as part of their responsibilities, the respondents are spread across four-year public institutions (64%), four-year private institutions (24%), and other institution types (13%).

Table 1.1. Demographic Characteristics of Teacher Survey Sample

Variable	Mean/ Proportion	SD	Min.	Max.	N
U.S. (1) vs. Int'l (0)	0.86		0	1	475
Age	50.04	11.65	25	77	337
Female	0.72		0	1	343
Assistant Prof.	0.11		0	1	345
Associate Prof.	0.20		0	1	345
Full Prof.	0.17		0	1	345
Full-time Non-TT	0.12		0	1	345
Administrator	0.07		0	1	345
Graduate Student	0.07		0	1	345
Other Position/NA	0.26		0	1	345
4-Year Public Institution	0.63		0	1	344
4-Year Private Institution	0.24		0	1	344
Other Institution	0.13		0	1	344
Teaches Composition	0.82		0	1	345
# of Students Graded per Sem.	3.33	1.12	1	8	344

Teachers also report substantially different grading workloads: the mean value of the categorical workload variable is 3.33 on a scale from 1-8, while the mode is 3 (a category that reflects a grading load of between 25 and 49 students per semester). As reflected in Figure 1.1, this distribution is skewed and has a high variance.

A significant number of respondents report direct grading responsibility for between 25 and 50 students per year. However, one non-negligible group of respondents reports having no grading responsibilities at all (perhaps because they have administrative roles that release them from teaching) while another group reports grading responsibilities that exceed 100 students per semester. Given that our survey has achieved substantial variation on this potentially important predictor of perceptions of peer feedback, we next proceed to developing and analyzing models that predict these attitudes.
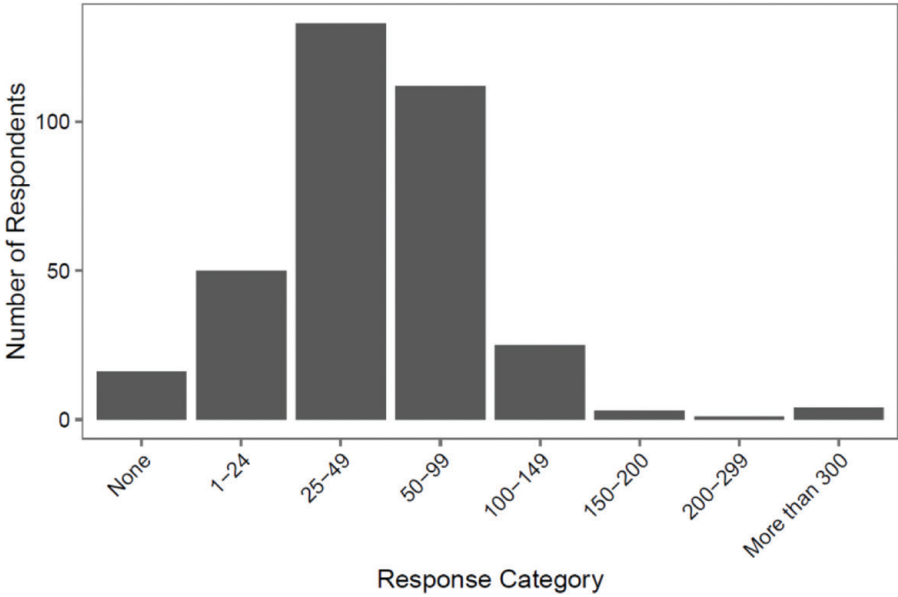
Figure 1.1 Distribution of respondents, grading load per semester (self-report).

RESULTS: PERCEPTIONS OF PEER FEEDBACK PRACTICES

First we examine the overall distribution of our sample's perceptions of instructor and peer feedback, as seen in Table 1.2. The results show that while few respondents perceive instructor or peer feedback to be of low overall helpfulness (3.2% and 0.3%, respectively), respondents are far less confident about instructor feedback than peer feedback. Only 12.8% of respondents perceive instructors to provide feedback of very high or high quality, compared to 25.8% stating that instructor feedback is likely to be of low or very low quality.

Table 1.2. Expert Perceptions of Instructor and Peer Feedback

Instructors Provide Feedback of . . .	Very High Quality	High Quality	Moderate Quality	Low Quality	Very Low Quality
Percent of Sample	0.9	11.5	61.8	22.6	3.2
Peer Feedback Is . . .	Extremely Helpful	Fairly Helpful	Moderately Helpful	Only a Little Helpful	Not at all Helpful
Percent of Sample	29.4	34.7	22	14.4	0.3

Note: Percentages may not sum to 100 due to rounding.

However, peer feedback is seen as "helpful" or "extremely helpful" by a narrow majority of respondents (64.1%), indicating teachers' optimism about the potential for this mode of feedback to improve the experience of novice writers. We can't probe the reasons why some respondents identify peer feedback as being "moderately," "only a little," or "not at all" helpful (36.7%), but this proportion of respondents is large enough to indicate that enthusiasm about peer feedback is by no means universal among writing teachers. What might predispose some teachers to view this feedback as helpful or unhelpful to students?

DETERMINANTS OF PEER FEEDBACK PERCEPTIONS

Earlier, we suggested that support for peer feedback may depend on considerations like grading load and faculty rank and age. In a linear regression model, we regress demographics and perceptions of instructor feedback on support for peer feedback practices. The results are presented in Table 1.3.

The results of Table 1.3 provide evidence that, as expected, demographics like gender and nationality have little impact on support for peer review practices, but institutional type, academic rank, and grading load each exert substantial effects on these perceptions. Notably, we find that respondents with academic appointments at private 4-year institutions are substantially less likely to support peer review (a decline of nearly 0.5 points on the 5-point scale; p = 0.002). This finding is interesting in light of assumptions that private institutions, especially smaller liberal arts colleges, emphasize the undergraduate learning experience, leading to more frequent use of experiential learning techniques. While the reason for the finding is not clear, it's possible that faculty at such institutions successfully use a greater number of other evidence-based techniques for writing instruction and therefore rate peer review as less useful relative to these strategies. For example, the lower student-teacher ratios at smaller liberal arts colleges may allow teachers to provide their own response on drafts instead of using peer review. Our data bear out this assumption: among respondents at private 4-year colleges, the mean on the teaching load variable is 3.125; among those at public 4-year colleges, the mean is 3.390 (p(t) = 0.023). Or it could be that these faculty are more likely to have experimented with peer review relative to other groups and found these experiences to be unsatisfactory. An alternative possibility is that faculty at these institutions are less likely than those at larger public four-year universities to have invested time to fully examine the current research on peer review or to have been introduced to the method and prepared to use it. Further studies could examine the less robust relationship between private college/university settings and support for peer review.

Table 1.3. Linear Regression Model Predicting Support for Peer Feedback Practices

Age	0.003
	(0.005)
Female	-0.069
	(0.127)
United States	-0.146
	(0.167)
Private 4-Year Inst.	-0.444***
	(0.142)
2-Year Inst.	-0.076
	(0.257)
Full Professor	-0.230
	(0.176)
Associate Prof.	0.310**
	(0.155)
Assistant Prof.	0.388**
	(0.180)
Administrator	0.142
	(0.240)
Teaches Composition	0.002
	(0.151)
Grading Load	-0.141***
	(0.052)
College-Level Feedback	-0.080
is High Qual.	(0.081)
Constant	4.586***
	(0.429)
Observations	328
R^2	0.084
Adjusted R^2	0.049
Residual Std. Error	1.003 (df = 315)
F Statistic	2.405*** (df = 12; 315)

Note: *p < 0.05 **p < 0.01 ***p < 0.001

While other findings from Table 1.3 provide evidence to support our expectations, one result runs counter to conventional wisdom. Assistant professors (around 0.3 points, $p = 0.047$) and associate professors (around 0.4 points, $p = 0.032$) have higher average perceptions of peer review than other groups, all else equal, which accords with the assumption that recently-minted and mid-career professors in the field have been exposed to current literature or teacher-development efforts supporting the use of peer feedback in the classroom. However, the unexpected finding to emerge from Table 1.3 is that a one-unit *increase* in grading load leads to a *decrease* in support for peer review of roughly 0.15 points ($p = 0.007$). Although this finding deserves more study, two possible reasons arise. First, because of their higher teaching loads, these individuals may be less likely to have been exposed to literature on peer feedback or have engaged in local faculty development efforts in the context of the time they must spend on their teaching. If the above expectation were to find support in future studies, it would point to the importance of lower class sizes and a stronger provision of workshops, training sessions, and the dissemination of best practices to faculty who are underexposed to state-of-the-art research because of pressures and responsibilities related to teaching, grading, and administration. Second, the additional time and effort required to manage and account for peer review (tracking exchanged papers, ensuring adequate response or evaluating its quality, comparing drafts and revisions, etc.) could be a disincentive for heavily burdened teachers who would rather use an "assign/collect/grade" approach to writing instruction.

FACULTY CHARACTERIZATION OF FEEDBACK

Our second research question was designed to explore teachers' conceptions of response quality based on the kind of language they expect to be used in high-quality responses and the kind of language they believe students use in peer review. In part, we wanted to see whether skepticism for peer review arises from a concern that students don't know how to provide high-quality responses based on the focus of their comments. If teachers believe that students provide responses similar to that provided by teachers, then skepticism must come from some other factor(s) than student ability.

To analyze the data we collected, we constructed document-term matrices that tabulate the presence of terms in each lexicon by respondent and term. These matrices create corpora of terms representing feedback to writing likely to be given by teachers and by students. In Figure 1.2, we manipulate these matrices to produce frequency histograms of the most common terms to appear in the dataset. This figure shows all terms that were mentioned by at least 10% of the respondents in the sample.
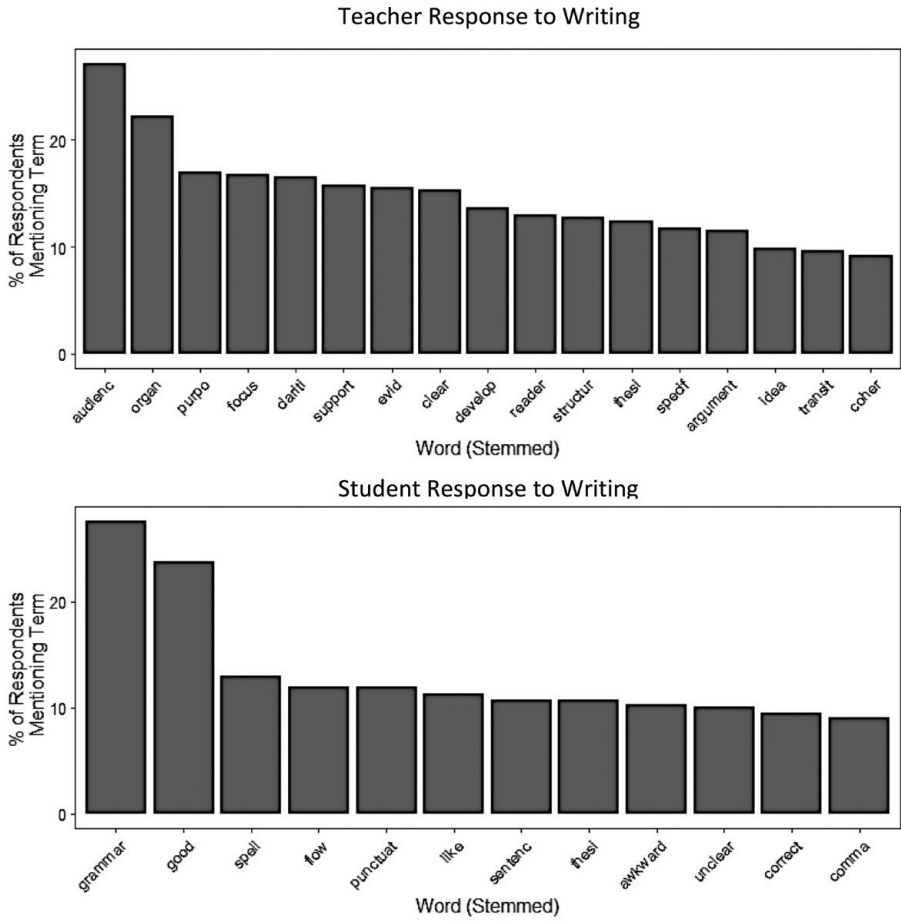
Teacher Response to Writing

Student Response to Writing

*Figure 1.2. Histograms of most frequent terms, teachers'
descriptions of teacher and student response to writing.*

Figure 1.2 shows that the teacher and student feedback lexicons differ substantially in several ways. First is the observation that teacher feedback contains more global, rhetorical, and conceptual-level terms like *audience*, *purpose*, *focus*, and *reader*. In comparison to these broad conceptual terms, teachers' expectations of student feedback include affective generalities (*good*, *like*, *awkward*) and sentence-level minutiae (*grammar*, *spelling*, *sentence*).

The teacher dataset also incorporates a greater overall number of terms related to evidence and support for arguments, indicating that this may be a primary concern among teachers with experience responding to developing writers. Interestingly, teachers are also more often in agreement about the key terms: a greater number of terms are *shared* by at least 10% of the respondents in the

expert lexicon. When it comes to student response, fewer teachers could agree on the most relevant keywords—there is less consensus about what kinds of concepts are likely to emerge in student response to writing. Perhaps this heterogeneity is related to respondents' different grading loads, meaning that some respondents have had fewer opportunities to gain a working understanding of how student writers approach revision. Whatever the reason, this pattern holds despite a greater diversity of teacher-oriented terms mentioned across all respondents in the dataset (after data cleaning, 964 unique teacher-oriented terms were collected, compared to 749 student-oriented terms).

SOURCES OF VARIATION IN TEACHER LEXICONS

While the above description shows some meaningful variation in the way that teachers characterize student and teacher feedback, we also examined how the demographic characteristics and perceptions of the respondents influenced the patterns of keyword mentions. That is, did the response lexicons vary as a function of workload, academic position, and/or other variables? Our first glimpse at these relationships comes from Table 1.4, in which we present binary logistic regression models predicting the likelihood of term occurrence. Each column in Table 1.4 represents a separate regression model, in which predictors of the most prevalent teacher-associated feedback terms include age, gender, nationality, type of institution, academic rank, grading load, and perceptions about feedback.

The results of Table 1.4 show that while some variables like institution type, composition instructor, and administrator role exert minimal influence on the incidence of terms across the dataset, other variables, most notably institution, have more meaningful effects. Across the terms in question, respondents from outside the United States (likely those who responded from the EATAW listserv) provide very different types of feedback. This may reflect some differences in translation, despite the survey being conducted in English; it may also be rooted in cultural and educational differences in feedback practices and the "language of response" in different countries, which is a subject deserving further study.

It also appears that for terms such as *audience* and *purpose*, which reflect knowledge of contemporary rhetorical approaches to writing, one of the most influential variables is instructor feedback perceptions. On average, a teacher who becomes one unit more positive in their perceptions of instructor feedback in higher education is expected to be around 1.7 times more likely to mention "audience," for example ($p < 0.01$). This substantial difference contrasts many nonsignificant predictors of these key terms: it appears that teachers who are optimistic about the quality of instructor feedback, regardless of academic position or demographics, are more likely to associate these key concepts with high-quality response.

Table 1.4. Binary Logistic Regression Predicting Likelihood of Use of "Principled" Terms (Part 1)

	Dependent variable:				
	Audience	Organization	Purpose	Focus	Clarity
Age	-0.020*	-0.002	0.018	0.008	-0.0004
	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)
Female	0.496*	0.152	-0.029	-0.359	1.032***
	(0.283)	(0.284)	(0.302)	(0.296)	(0.374)
United States Institution	1.904***	1.963***	2.057***	2.618**	-0.108
	(0.518)	(0.558)	(0.756)	(1.034)	(0.386)
Private 4-Year Inst.	-0.364	-0.215	-0.300	0.222	0.059
	(0.311)	(0.320)	(0.355)	(0.331)	(0.348)
2-Year Inst.	-0.438	0.766	0.276	0.116	-0.109
	(0.588)	(0.543)	(0.591)	(0.598)	(0.693)
Full Professor	0.228	-0.938**	-0.011	0.160	-0.419
	(0.390)	(0.416)	(0.423)	(0.434)	(0.448)
Associate Prof.	0.703**	-0.164	0.432	0.870**	-0.156
	(0.339)	(0.336)	(0.368)	(0.370)	(0.368)
Assistant Prof.	0.210	-0.341	0.150	0.434	-0.994*
	(0.400)	(0.412)	(0.464)	(0.462)	(0.529)
Administrator	0.047	-0.118	-0.110	0.662	0.053
	(0.516)	(0.503)	(0.592)	(0.543)	(0.546)
Teaches Composition	-0.138	-0.056	-0.510	0.114	0.206
	(0.350)	(0.363)	(0.438)	(0.387)	(0.353)
Grading Load	-0.122	-0.037	-0.202	0.023	0.090
	(0.122)	(0.125)	(0.150)	(0.144)	(0.120)
Perception of Feedback Quality	0.534***	0.007	0.421**	0.087	-0.239
	(0.187)	(0.185)	(0.207)	(0.209)	(0.197)
Constant	-3.282***	-2.244**	-3.844***	-4.586***	-1.544
	(1.064)	(1.087)	(1.318)	(1.496)	(1.040)
Observations	329	329	329	329	329
Log Likelihood	-192.410	-187.448	-161.038	-161.127	-165.817
Akaike Inf. Crit.	410.821	400.897	348.075	348.255	357.635

*Note: *p<0.1; **p<0.05; ***p<0.01*

Table 1.4. Binary Logistic Regression Predicting Likelihood of Use of "Principled" Terms (Part 2)

	Dependent variable:				
	Support	Evidence	Clear	Develop	Reader
Age	0.002	-0.001	-0.011	0.003	0.033**
	(0.013)	(0.013)	(0.013)	(0.015)	(0.016)
Female	-0.012	0.583*	-0.117	0.146	0.240
	(0.308)	(0.342)	(0.318)	(0.341)	(0.389)
United States Institution	1.348**	1.350**	0.150	1.568**	0.379
	(0.632)	(0.570)	(0.428)	(0.762)	(0.522)
Private 4-Year Inst.	0.229	-0.061	0.085	0.050	-0.644
	(0.336)	(0.349)	(0.355)	(0.386)	(0.454)
2-Year Inst.	0.225	-0.732	0.541	0.116	-0.834
	(0.590)	(0.797)	(0.629)	(0.633)	(1.085)
Full Professor	0.458	0.288	-0.196	0.388	-0.683
	(0.421)	(0.425)	(0.436)	(0.467)	(0.553)
Associate Prof.	0.310	0.101	-0.603	0.277	0.533
	(0.385)	(0.383)	(0.421)	(0.417)	(0.416)
Assistant Prof.	0.093	-0.209	-0.293	0.138	-0.511
	(0.476)	(0.488)	(0.460)	(0.508)	(0.665)
Administrator	0.696	-0.727	-0.154	0.342	-0.378
	(0.539)	(0.679)	(0.579)	(0.642)	(0.711)
Teaches Composition	-0.148	0.349	0.682*	-1.101*	0.330
	(0.393)	(0.372)	(0.350)	(0.566)	(0.439)
Grading Load	0.031	-0.057	-0.173	0.077	-0.223
	(0.140)	(0.142)	(0.135)	(0.156)	(0.160)
Perception of Feedback	-0.165	0.063	-0.033	0.325	0.164
Quality	(0.212)	(0.209)	(0.204)	(0.231)	(0.231)
Constant	-2.315*	-3.385***	-1.090	-3.521**	-3.432***
	(1.220)	(1.202)	(1.048)	(1.473)	(1.305)
Observations	329	329	329	329	329
Log Likelihood	-161.921	-157.705	-160.238	-138.611	-123.312
Akaike Inf. Crit.	349.841	341.409	346.476	303.222	272.624

*Note: *p<0.1; **p<0.05; ***p<0.01*

Table 1.5 shows differences in teachers' mention of student terms across key demographic and perceptual measures. Many respondents seemed to have fewer things to say overall about student feedback, perhaps reflecting some respondents' lack of experience seeing what students write to each other. However, several patterns emerged. While the country of origin of the institution again plays a major role in predicting the lexicon used to describe student response, another variable of interest is that of institutional type. Respondents who identified as instructors at 2-year colleges were more likely than average to mention terms corresponding to affective considerations, such as "like" ($p < 0.01$) and "unclear" ($p < 0.05$), but less likely to mention "grammar," perhaps reflecting a belief (or experience) that their students more often identify surface problems experientially ("I'm confused") than concerning explicit rules ("this comma splice obscures the meaning of your sentence"). However, across the survey, respondents actively teaching composition were less likely than average to mention the affective term "good," perhaps because they more systematically direct their students toward specifics. These and other patterns relating to the student lexicon reflect the heterogeneity that characterizes teachers' determination or understanding of student feedback.

A CLOSER LOOK AT PEER FEEDBACK PERCEPTIONS

In the preceding analyses, we examined differences in the teacher lexicons on the basis of demographics and perceptions of overall feedback quality. Now we take a closer look at perceptions of *peer feedback* as a critical determinant of the type of content used by respondents. This analysis allows us to examine, consistent with the findings above, whether skeptics of student peer review think about feedback differently than their more supportive counterparts.

To perform the analysis, we divided the sample into those expressing positive views of peer feedback (it is "extremely helpful" or "helpful") and those expressing neutral or pessimistic views. The distribution of the resulting binary variable leaves us with a tally of 36.4% skeptics and 63.6% proponents in the sample. Figure 1.3 provides a depiction of *differences* in the likelihood that these two groups mention the most prevalent keywords in the "high-quality" (teacher) corpus. The left-hand side of the figure shows terms that skeptics were substantially more likely to mention than supporters of peer feedback (black bars denote statistical significance on the basis of Welch two-sample t-tests at the $p < 0.05$ level). On the right-hand side of the figure, we see the opposite: these are terms that *proponents* of peer feedback mentioned substantially more than skeptics. Terms in the middle were mentioned by both proponents of peer feedback and skeptics at roughly equivalent rates.

Table 1.5. Binary Logistic Regression Predicting Likelihood of Mention of Student Terms (Part 1)

	Dependent variable:				
	Grammar	Good	Spell	Flow	Punctuation
	(1)	(2)	(3)	(4)	(5)
Age	0.002	-0.017	0.016	-0.004	0.041***
	(0.011)	(0.012)	(0.014)	(0.015)	(0.015)
Female	-0.081	0.115	0.214	0.517	0.163
	(0.267)	(0.283)	(0.340)	(0.380)	(0.349)
United States Institution	0.693*	0.589	0.235	1.784**	0.436
	(0.392)	(0.404)	(0.456)	(0.781)	(0.510)
Private 4-Year Inst.	0.220	-0.200	-0.598	0.044	-0.294
	(0.294)	(0.326)	(0.403)	(0.395)	(0.392)
2-Year Inst.	-1.146*	0.165	-0.254	0.477	-0.692
	(0.678)	(0.534)	(0.692)	(0.713)	(0.816)
Full Professor	-0.055	0.071	-0.231	-0.403	-0.227
	(0.375)	(0.394)	(0.496)	(0.524)	(0.506)
Associate Prof.	0.695**	0.017	0.610	0.110	0.832**
	(0.323)	(0.341)	(0.390)	(0.416)	(0.405)
Assistant Prof.	-0.169	-0.125	0.546	0.038	0.703
	(0.395)	(0.398)	(0.461)	(0.505)	(0.496)
Administrator	-0.141	-0.561	0.443	-1.075	0.578
	(0.511)	(0.607)	(0.590)	(0.829)	(0.605)
Teaches Composition	-0.537	-0.822**	-0.070	-0.896	-0.217
	(0.340)	(0.401)	(0.419)	(0.571)	(0.443)
Grading Load	-0.196*	0.095	0.024	-0.423**	0.037
	(0.119)	(0.119)	(0.133)	(0.187)	(0.143)
Feedback Qual. Percep.	-0.074	0.124	0.019	0.152	0.106
	(0.174)	(0.181)	(0.209)	(0.239)	(0.219)
Constant	0.242	-0.637	-2.545**	-1.706	-3.907***
	(0.943)	(1.013)	(1.173)	(1.446)	(1.290)
Observations	329	329	329	329	329
Log Likelihood	-204.539	-191.237	-150.745	-128.048	-140.132
Akaike Inf. Crit.	435.078	408.474	327.489	282.095	306.265

*Note: *p<0.1; **p<0.05; ***p<0.01*

Table 1.5. Binary Logistic Regression Predicting Likelihood of Mention of Student Terms (Part 2)

|	Dependent variable				
	Like	Sentence	Thesis	Awkward	Unclear
	(6)	(7)	(8)	(9)	(10)
Age	0.025	0.018	0.014	-0.012	-0.018
	(0.016)	(0.015)	(0.015)	(0.015)	(0.016)
Female	0.033	-0.104	-0.240	0.047	0.988**
	(0.361)	(0.363)	(0.346)	(0.367)	(0.472)
United States Institution	1.572**	0.298	1.416*	0.350	-0.549
	(0.770)	(0.538)	(0.765)	(0.509)	(0.439)
Private 4-Year Inst.	-0.795	0.343	-0.196	-0.203	-0.411
	(0.493)	(0.391)	(0.423)	(0.422)	(0.471)
2-Year Inst.	1.619***	-0.436	-0.159	-0.996	1.108*
	(0.581)	(0.815)	(0.698)	(1.071)	(0.672)
Full Professor	-0.300	-0.077	-0.030	-0.044	-0.570
	(0.513)	(0.508)	(0.500)	(0.511)	(0.644)
Associate Prof.	0.410	0.514	0.155	-0.537	0.320
	(0.432)	(0.415)	(0.437)	(0.487)	(0.447)
Assistant Prof.	0.555	-0.070	0.295	-0.004	0.165
	(0.516)	(0.560)	(0.513)	(0.492)	(0.506)
Administrator	-1.320	-1.169	0.026	0.129	0.142
	(1.090)	(1.075)	(0.715)	(0.636)	(0.707)
Teaches Composition	-0.806	-0.402	-1.585**	-0.219	0.106
	(0.648)	(0.490)	(0.755)	(0.457)	(0.451)
Grading Load	-0.207	0.064	0.017	-0.115	-0.110
	(0.185)	(0.160)	(0.169)	(0.152)	(0.146)
Feedback Qual. Percep.	0.407*	0.025	0.171	0.333	0.151
	(0.243)	(0.236)	(0.240)	(0.232)	(0.233)
Constant	-3.816**	-2.580*	-2.245	-2.012	-1.905
	(1.586)	(1.358)	(1.585)	(1.239)	(1.243)
Observations	329	329	329	329	329
Log Likelihood	-120.323	-126.683	-126.551	-130.979	-123.105
Akaike Inf. Crit.	266.647	279.365	279.102	287.959	272.210

*Note: *p<0.1; **p<0.05; ***p<0.01*
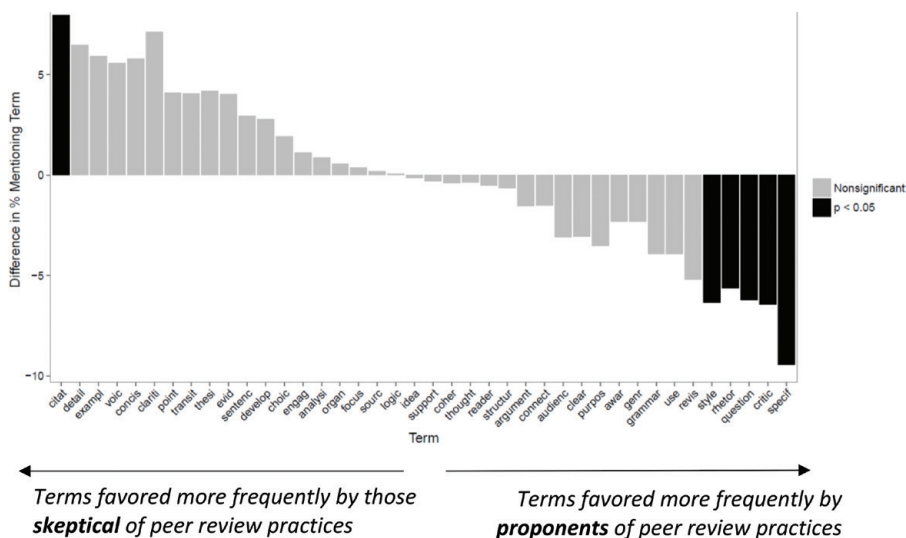
Figure 1.3. Comparison of high-quality term mention rate,
peer review skeptics vs. proponents of peer review

The results presented in Figure 1.3 demonstrate interesting patterns among proponents and skeptics of peer review. The leftmost bars show that many key terms are mentioned substantially more often by skeptics, despite the fact that only "citation" attains statistical significance at the $p < 0.05$ level. However, these terms, like "detail," "example," "concision," and "point," suggest that skeptics of peer review devote more attention to sentence-level concerns than do proponents. In fact, we see that peer feedback skeptics are around 35% more likely than proponents of peer review to mention the word "sentence" (a difference in the rate of mention of roughly 3%), though this difference is not statistically differentiable from 0 ($t = 0.87$, $p = 0.39$).

The rightmost bars of Figure 1.3, however, reveal more statistically significant differences when considering terms that skeptics used *less* often than proponents of peer review. Here, many broader concepts like "question," "criticism/critique," "style," and "revision" are used more frequently by proponents of peer review than by skeptics. To a lesser (nonsignificant) extent, we also see that rhetorical terms common in contemporary approaches to writing, such as "audience," "purpose," and "genre," are used more frequently by peer feedback proponents.

Taken together, these results point to differences in the way that skeptics and proponents of peer review think about high-quality feedback. It may be that this relationship occurs because those with a greater focus on specifics and mechanics in writing find peer review to be a dubious method to help students improve

their writing: if the purpose of peer review is to provide broad structural, rhetorical, and informational responses, there may be more trust in students' abilities than if the purpose is for students to find errors, for which they may not have appropriate skills (Anson, "Talking"). Or it could be the reverse: teachers who have had negative experiences with peer review practices might find themselves increasingly prioritizing surface-level matters (such as error) in their own responses, as they find them to be critically overlooked across students' evaluations of writing. Regardless, future studies should further investigate the causal roots of this relationship.

CONCLUSION: THE ROAD AHEAD FOR RESEARCH AND INSTRUCTION IN PEER REVIEW

Results of our analysis demonstrate that while overall, teachers of writing appear fairly receptive to the idea of peer review, considerable variation exists across public and private university settings, as well as across academic ranks (though age does not play a role). Interestingly, we also observe decreasing support for peer review practices among instructors tasked with heavier grading loads. Analyses of key term usage also show differences across national context and perceptions of peer feedback effectiveness. We also find that instructors who are pessimistic about the implementation of peer review identify different concepts as important components of teacher-provided response compared to the response provided by students.

Taken together, these results suggest that the field has asymmetrically incorporated peer review in writing instruction. We must continue to advocate for the practice of peer review, which the pedagogical literature as well as newer research supports on the basis not only of improved final texts but practice of revision and the learning of useful collaborative communicative skills often expected in the workforce (Bruffee). Advocates of peer review argue that it is most effective when instructors fully orient students to the process and coach them in how to provide insightful feedback. These orientations include working through a sample draft together with the class, showing videos of successful response sessions and ones that get derailed for different reasons, providing peer-response guides that help students to focus on specific issues of importance to the development of their drafts, and asking for meta-commentaries of the peer-review sessions after they're done. Others focus their advice on ways to incorporate good peer review practices within the context of a well-supported and integrated approach to writing at the department or institutional level (Anson, Gonyea, and Paine). These approaches might involve leveraging new technologies adapted to such a task, such as digitally-mediated peer review systems (Moxley) or calibrated peer

review programs (Reynolds and Moskovitz). These and other approaches may involve providing greater support for faculty hoping to incorporate peer review.

Further research is also needed to study the relationship between teachers' and students' attitudes toward peer review (for the latter, see Mulder, Pearce, and Baik), because this relationship may also influence how effectively students use the process. Deeper and more robust information is needed about teachers' opinions of peer review based on their experiences; variations in the use of peer review as these relate to its success; the experiences of students as they engage in peer review; and peer review as a function of assignments, genres, learning contexts, developmental stages of students as writers, instructor variables such as ideologies of teaching and learning, and student variables such as measures of self-efficacy, writing anxiety, and prior experience.

Finally, and perhaps most importantly, scholarship on threshold concepts is increasingly pointing to the relationship between the language of writing instruction and the underlying concepts and understandings associated with the production and use of written text and the ability of writers to "transfer" their understandings to other contexts and genres (Adler-Kassner and Wardle; Anson and Moore; Yancey, Robertson, and Taczak). Downs and Robertson, for example, suggest four domains of threshold concepts to emphasize in foundational writing courses: human interaction (rhetoric); textuality; epistemology (ways of knowing); and process. Comparing skilled and novice writers, they point to differences directly relevant to peer review:

> Seasoned writers usually treat writing as a rhetorical human interaction in which readers and writers interact to shape writing and meaning. Novice writers are much less likely to recognize the interactional nature of writing. To them, writing is strictly about getting sentences right rather than interacting with or being responsible to readers. Building an understanding of writing as a rhetorical activity, as human interaction, seems an essential threshold concepts for FYC [first-year composition]. 107

Because an understanding of such threshold concepts is revealed in the language and terminology writers use to talk about their own and others' writing, peer review will succeed or fail in proportion to this understanding. Reciprocally moving between the experience and practice of peer review and discussion of the meta-level threshold concepts at the heart of successful writing may strengthen students' abilities to respond to each other's writing and subsequently build confidence in teachers that the time spent in peer review will help students to grow as writers.

WORKS CITED

Adler-Kassner, Linda, and Elizabeth Wardle, eds. *Naming What We Know: Threshold Concepts of Writing Studies*. Utah State UP, 2015.

Anderson, Paul, et al. "The Contributions of Writing to Learning and Development: Results from a Large-Scale Multi-institutional Study." *Research in the Teaching of English*, vol. 50, no. 2, 2015, pp. 199-235.

Anson, Chris M. *Talking About Mistakes: Mining Students' Oral Peer Reviews for Constructs and Practices in Error Detection.* Paper delivered at the European Association for Research on Language and Instruction (EARLI) SIG Writing conference, Liverpool, England, July 6, 2016.

---. "Process Pedagogy and Its Legacy." *A Guide to Composition Pedagogies* (2nd ed.), edited by Amy Rupiper Taggart et al., Oxford UP, 2013, pp. 212-30.

---. "Response and the Social Construction of Error." *Assessing Writing,* vol. 7, no. 1, 2000, pp. 5-21.

Anson, Chris, ed. *Writing and Response: Theory, Practice, Research*. National Council of Teachers of English, 1989.

Anson, Chris M., and Jessie L. Moore, editors. *Critical Transitions: Writing and the Question of Transfer*. The WAC Clearinghouse and UP of Colorado, 2016.

Anson, Ian G., and Chris M. Anson. "Assessing Peer and Instructor Response to Writing: A Corpus Analysis from an Expert Survey." *Assessing Writing*, Vol. 33, 2017, pp. 12-24.

Anson, Chris M., et al. "Considering What They Say: An Analysis of Key Terms Used by Faculty Across the Curriculum." *Considering What We Know: Threshold Concepts for Writing Studies*, edited by Linda Adler-Kassner and Elizabeth Wardle, Utah State UP, 2019, pp. 314-27.

Bailey, Richard, and Mark Garner. "Is the Feedback in Higher Education Assessment Worth the Paper it is Written On? Teachers' Reflections on their Practices." *Teaching in Higher Education*, vol. 15, no. 2, 2010, pp. 187-98.

Bean, John C. *Engaging Ideas: The Professor's Guide to Integrating Writing, Critical Thinking, and Active Learning in the Classroom.* 2nd ed., Jossey-Bass, 2011.

Braine, George. *Writing Across the Curriculum: A Case Study of Faculty Practices at a Research University*. National Center for Research on Teacher Learning, 1990. (*ERIC* Document Reproduction Service ED 324 680)

Brammer, Charlotte, and Mary Rees. "Peer Review from the Students' Perspective: Invaluable or Invalid?" *Composition Studies,* vol. 35, no. 2, 2007, pp. 71-85.

Breuch, Lee-Ann Kassman. *Virtual Peer Review: Teaching and Learning About Writing in Online Environments*. SUNY Press, 2012.

Bruffee, Kenneth A. "Learning and the 'Conversations of Mankind.'" *College English, vol.* 46, no. 7, 1984, pp. 635-52.

Cho, Kwangsu, et al. "Commenting on Writing: Typology and Perceived Helpfulness of Comments from Novice Peer Reviewers and Subject Matter Experts." *Written Communication*, vol. 23, no. 3, 2006, pp. 260-94.

Downs, Doug, and Liane Robertson. "Threshold Concepts in First-Year Composition. *Naming What We Know: Threshold Concepts for Writing Studies*, edited by Linda Adler-Kassner and Elizabeth Wardle, Utah State UP, 2015, pp. 105-21.

Elbow, Peter. "Closing My Eyes as I Speak: An Argument for Ignoring Audience." *College English,* vol. 49, no. 1, 1987, pp. 50-69.

Elbow, Peter. *Writing Without Teachers.* Oxford UP, 1973.

Elbow, Peter, and Pat Belanoff. *Sharing and Responding*. Random House, 1989.

Falchikov, Nancy, and Judy Goldfinch. "Student Peer Assessment in Higher Education: A Meta-analysis Comparing Peer and Teacher Marks. *Review of Educational Research,* vol. 70, 2000, pp. 287-322.

George, Diana. "Writing with Peer Groups in the Composition Classroom." *College Composition and Communication,* vol. 35, no. 3, 1984, pp. 320-26.

Gere, Anne Ruggles. *Writing Groups: History, Theory, and Implications*. Southern Illinois UP, 1987.

Graff, Nelson. "Approaching Authentic Peer Review." *The English Journal,* vol. 98, no. 5, 2009, pp. 81-87.

Haswell, Richard. "NCTE/CCCC's Recent War on Scholarship." *Written Communication*, vol. 22, no. 2, 2005, pp. 198-223.

Howard, Rebecca Moore. "Collaborative Pedagogy." *Composition Pedagogies: A Bibliographic Guide*, edited by Gary Tate, Amy Rupiper, and Kurt Schick. Oxford UP, 2001, pp. 54-70.

Illich, Lindsay. "How to See a Text: The Word Cloud Peer Review." *Journal of Teaching Writing*, vol. 28, no. 2, 2013, pp. 1-18.

Knoblauch, Cy H., and Lil Brannon. "Teacher Commentary on Student Writing: The State of the Art." *Freshman English News,* vol. 10, 1981, pp. 1-4. Rpt. in Richard Graves, ed., *Rhetoric and Composition: A Sourcebook for Teachers and Writers*. Boynton/Cook, 1984.

Jesnek, Lindsey M. "Peer Editing in the 21st Century College Classroom: Do Beginning Composition Students Truly Reap the Benefits?" *Journal of College Teaching and Learning,* vol. 8, no. 5, 2011, pp. 17-24.

Leijen, Djuddah. "Applying Machine Learning Techniques to Investigate the Influence of Peer Feedback on the Writing Process." *Methods in Writing Process Research*, edited by D. Knorr et al. Peter Lang Verlag, 2014, pp. 167-83.

Li, Jinrui, and Roger Barnard. "Academic Tutors' Beliefs about and Practices of Giving Feedback on Students' Written Assignments: A New Zealand Case Study." *Assessing Writing*, vol. 16, no. 2, 2011, pp. 137-48.

Moxley, Joseph. "Datagogies, Writing Spaces, and the Age of Peer Production." *Computers and Composition*, vol. 25, no. 2, 2008, pp. 182-202.

Moxley, Joseph M., and David Eubanks. "On Keeping Score: Instructors' vs. Students' Rubric Ratings of 46,689 Essays." *WPA: Writing Program Administration,* vol. 39, no. 2, 2016, pp. 53-80.

Mulder, Raoul A., et al. "Peer Review in Higher Education: Student Perceptions Before and After Participation." *Active Learning in Higher Education*, vol. 15, no. 2, 2014, pp. 157-70.

Nelson, Melissa M., and Christian D. Schunn. "The Nature of Feedback: How Different Types of Peer Feedback Affect Writing Performance." *Instructional Science*, vol. 37, no. 4, 2009, pp. 375-401.

Nicol, David. "From Monologue to Dialogue: Improving Written Feedback Processes in Mass Higher Education." *Assessment & Evaluation in Higher Education*, vol. 35, no. 5, 2010, pp. 501-17.

Nicol, David, et al. "Rethinking Feedback Practices in Higher Education: A Peer Review Perspective." *Assessment & Evaluation in Higher Education*, vol. 39, no. 1, 2014, pp. 102-22.

Paulson, Eric J., et al. "Peer Review Re-Viewed: Investigating the Juxtaposition of Composition Students' Eye Movements and Peer Review Processes." *Research in the Teaching of English,* Vol. 41, No. 3, 2007, pp. 304-35.

Reynolds, J., and Cary Moskovitz, "Calibrated Peer Review Assignments in Science Courses: Are They Designed to Promote Critical Thinking and Writing Skills?" *Journal of College Science Teaching,* vol. 38, no. 2, 2008, pp. 60-66.

Roskelly, Hephzibah. "The Risky Business of Group Work." *ATAC Forum 4*, Spring, 1992. Reprinted in *The Writing Teacher's Sourcebook,* edited by Gary Tate et al., 3rd ed., Oxford UP, 1994, pp. 141-48.

Sommers, Nancy. "Across the Drafts." *College Composition and Communication*, vol. 58, no. 2, 2006, pp. 248-56.

Sommers, Nancy. *Beyond the Red* Ink (video). St. Martins, 2012. Available at https://www.youtube.com/watch?v=PKfLRz7h7gs.

Sommers, Nancy. *Responding to Student Writin*g. Bedford/St. Martin's, 2013.

Sommers, Nancy. "Revision Strategies of Student Writers and Experienced Adult Writers." *College Composition and Communication,* vol. 31, no. 4, 1980, pp. 378-88.

Spear, Karen. *Shared Writing: Peer Response Groups in English Classes.* Boynton/Cook, 1988.

Stewart, Trevor Thomas, and Greg McClure. "Freire, Bakhtin, and Collaborative Pedagogy: A Dialogue with Students and Mentors." *International Journal of Dialogic Science,* vol. 7, no. 1, 2013, pp. 91-108.

Straub, Richard. *The Practice of Response: Strategies for Commenting on Student Writing*. Hampton Press, 2000.

Straub, Richard. "Responding—Really Responding—to Other Students' Writing." *The Subject is Writing,* edited by Wendy Bishop. Boynton/Cook, 1999, pp. 136-46.

Tuck, Jackie. "Feedback-Giving as Social Practice: Teachers' Perspectives on Feedback as Institutional Requirement, Work and Dialogue." *Teaching in Higher Education*, vol. 17, no. 2, 2012, pp. 209-22.

Yancey, Kathleen Blake, Liane Robertson, and Kara Taczak. *Writing Across Contexts: Transfer, Composition, and Site of Writing.* Utah State UP, 2015.