

14

WRITING TO LEARN AND LEARNING TO WRITE THROUGH SWoRD

Christian Schunn

Introduction

The largest barrier to giving students opportunities to work on rich problems and tasks requiring a free text response is a feedback bottleneck: there are so many students and there is so little time to provide students with detailed and useful feedback. New automated feedback tools can address a number of those situations, but are not general enough to be applied in all contexts. By contrast, peer feedback is a highly flexible strategy that can be applied to essentially any task assigned to students. Whenever students are given a writing task, they can also be asked to evaluate their peers' work using rubrics and provide constructive criticism. Thus, students can receive detailed feedback on short or long writing tasks without always requiring instructor feedback. Further, students can also learn from evaluating their peer's writing.

To make peer feedback a viable strategy, students need to be strongly supported in the process and held accountable for their work so they provide accurate and detailed feedback rather than short and superficial feedback (i.e., "good job"). SWoRD (Scaffolded Writing and Rewriting in the Discipline) is a tool that has been iteratively improved through research (described below) to give the support students need to effectively and efficiently provide feedback on the aspects of the assignment that are important for instruction. Most saliently, it has algorithms that ensure students take the reviewing task seriously. Teachers can then assign many more rich writing-based assignments than they normally would be able to without being overwhelmed by grading/feedback workload.

Writing as Central but Challenging

Writing is a powerful instructional tool. It can be used to support students' development of self-regulation skills, and self-regulation is a critical part of effective

writing itself (Dignath, Buettner, & Langfeldt, 2008; Kliewer et al., 2011). Across diverse disciplines, in written responses, students can show that they understand how and why their answers were obtained (Miller, 1992; Rubin & Carlan, 2005), rather than potentially having guessed a correct response on a multiple-choice exam.

Recent standards in mathematics and science are full of performance indicators that require writing, and writing is likewise required in corresponding assessments. For example, in mathematics, the Common Core standards ask that students “Understand solving equations as a process of reasoning and explain the reasoning.” Similarly, in science, the Next Generation Science Standards frequently refer to practices like “make and defend a claim,” or “explain.” Until students can turn in video answers or be interviewed one-on-one in assessments, evaluations of students’ ability to explain their reasoning will likely involve writing.

English Language Arts (ELA) has always required writing, and the new Common Core standards for ELA have placed even greater emphasis on higher-level writing skills (<http://www.corestandards.org/ELA-Literacy/>). Many of these standards refer to writing skills that are useful across other topic areas, such as history and science. For example, these standards require that students “Cite strong and thorough textual evidence to support analysis of what the text says explicitly.”

At the same time, writing is a very complex skill that students struggle to master. A main source of difficulty relates to the many layers of language that must be mastered to produce functional text. For instance, there are the complex lower levels of a language: (1) how letters form words (especially for languages like English that are filled with irregular spellings rather than languages like German); and (2) how words are sequenced to form grammatical sentences. In addition, there are the complex upper levels of language: (1) how sentences come together to form coherent arguments or narratives, building on what is already established in the text; and (2) how all of the text relates to what an intended audience member already knows about the topic. In other words, writing is an integration of many levels of skills, and developing writers need time to develop all of these skills.

Another source of difficulty comes from the integration of discipline content with these complex writing skills. Writing across disciplines takes on new forms or genres (Biber & Conrad, 2009) with each discipline potentially having their own conventions (e.g., a lab report in science). Students then have more forms to learn. But just as problematic, students are also struggling with the underlying disciplinary ideas about which they are writing (e.g., what is a phylum, how do laws shape a democracy, how do substances dissolve). This creates a fundamental working memory challenge for students (Hayes, Flower, Schriver, Stratman, & Carey, 1987; Kellogg, 1994) because of having to process complex disciplinary ideas, which are still just many isolated pieces, while also thinking about how to write text that is coherent and clear.

In sum, writing is cognitively difficult from the complexity of things to learn and from the high load on the student’s working memory. These challenges

are both addressed through carefully structured practice (Kellogg & Whitford, 2009), as students come to master all of the layers of text production in different genres and develop integrated conceptual understandings. Students need to be given opportunities to think through the various aspects and practice each aspect (see also Crossley, Allen, & McNamara, Chapter 12 this volume).

The Need for Effective Peer Feedback

If such carefully structured practice with writing tasks is the answer, ironically, practice (of any kind) is what is currently most lacking. Students rarely write in general, and they especially do not very often write anything long enough to involve a real argument or an interesting narrative (Kihara, Graham, & Hawken, 2009). For example, a national study of writing practice in middle and high schools in the US found that students do very little extended writing in English class, and almost no extended writing at all in science, social studies, or math classes (Applebee & Langer, 2011). Further, students tend not to receive timely or effective feedback from teachers even when they do write.

As a result, student writing performance in the US is quite poor by national assessments, with most 8th and 12th graders writing at the basic or below basic levels (NAEP, 2011), and little improvement has been shown over the last 20 years (National Commission on Writing in American Schools and Colleges, 2003). Students are not entering the workplace or college ready for what is expected there (National Commission on Writing, 2004).

Why do students not get opportunities to practice? The high workload associated with grading and feedback is a primary culprit. As one teacher said quite clearly, “Well, I can’t have them write two paragraphs every day because that will take me how much time to read and if I can’t read it and give them thoughtful feedback, it’s not very productive” (Applebee & Langer, 2011). The need for feedback for effective writing practice is obvious; what is needed is an alternative method for practice with feedback that is not dependent upon teacher feedback alone.

Automatic Essay Scoring and Automatic Writing Evaluation systems will be an important part of the solution to this lack of practice and feedback problem (see Crossley, Allen, & McNamara, Chapter 12 this volume; Ramineni & Deane, Chapter 10 this volume). However, there will likely be some forms of writing that are not easily handled using those automatic systems (e.g., ones involving integration of text and images, because those systems cannot parse images), or because of significant involvement of content understanding (e.g., distinguishing between the presence of a supporting clause and one that involves sensible content that actually is supporting).

Peer feedback can be applied to any kind of writing task assigned to students. The basic notion behind peer feedback is that if students can be reasonably expected to provide a response to the task, they can also be reasonably expected to be able

to evaluate strengths/weaknesses in other students' written responses and suggest possible improvements. This follows from a general developmental pattern in which students can recognize what is needed well before they can reliably do what is needed themselves (Siegler, 1996; Vygotsky, 1978). For example, a person can talk about and recognize good/bad driving from taking a driving course well before they are good drivers. As discussed below, peer ratings of writing on clear rubrics are generally very accurate.

More importantly, giving peer feedback is another excellent learning opportunity; by evaluating responses using a rubric and by providing constructive comments to the author, students can improve their own detection and revising skills (Palincsar & Brown, 1984; Topping, 2008; Van Den Berg, Admiraal, & Pilot, 2006). Further, it emphasizes writing as a process, rather than just writing as an outcome (Katstra, Tollefson, & Gilbert, 1987), which is an important transition in mindset about writing.

However, while peer feedback can and often is implemented in classrooms without technology support (Applebee & Langer, 2011), it is frequently difficult to orchestrate, consumes an unnecessarily large amount of precious class time, and often is poorly executed. Students struggle to provide honest evaluations, worrying about threatening friendships, confronting social power dynamics, or generally not wanting to embarrass the author in public. Also, students have little incentive to put effort into the evaluations. Thus, without additional support, students typically provide short, content-free positive evaluations such as "This is awesome, dude!" (VanDeWeghe, 2004).

The SWoRD System

There is a simple automated approach to creating effective student peer review in a wide range of classes, which has been iteratively improved over a decade of research (Cho & Schunn, 2007; Cho, Schunn, & Charney, 2006a; Cho, Schunn, & Wilson, 2006b; Kaufman & Schunn, 2011; Nelson & Schunn, 2009; Patchan, Hawk, Stevens, & Schunn, 2013). At its core, there are four elements of SWoRD:

1. a method for easily assigning anonymized documents to peers and returning reviews back to authors via the web;
2. a structured reviewing form that contains concrete suggestions for what kinds of comments are requested on specific dimensions and rating rubrics that have concrete anchors for each rating level that are also tied to specific dimensions;
3. student authors rate the helpfulness of the reviews they receive, and these helpfulness ratings (called back-evaluations) are used to compute a reviewing helpfulness grade for the reviewers. This forces students to take the task of giving comments seriously;

- all ratings produced by a given student are compared to the mean ratings produced by all the other student reviewers of the same documents. This is done across the multiple rating rubrics (e.g., three to eight rubrics in an assignment) multiplied by the number of peers' documents reviewed by the student for the given assignment (e.g., 4 documents x 5 rubrics = 20 ratings). If the ratings are similarly ordered (from relative weaknesses to relative strengths), then the student receives a high reviewing accuracy grade. If the ratings are very dissimilar (either by giving random ratings or all the same ratings), then the student receives a low reviewing accuracy grade. This forces students to take the rating task seriously.

These four elements work together to produce a method that is easily integrated into diverse assignments and classes to produce accurate ratings and useful comments, and ultimately good learning outcomes. In early years, SWoRD was most commonly used in college settings, as the internet was more broadly available to college-age students. In recent years, web-based methods are now much easier to use in K-12 settings, as both schools and children's homes have achieved very broad access. With growth in use at the high school level, adjustments have been made to SWoRD that support the needs of younger learners and the various obligations that high school teachers more commonly face.

The Student Tasks: Submission, Reviewing, Back-Evaluation

Students complete three basic tasks for a given assignment draft, which are clearly shown to students in a timeline view (see Figure 14.1). There is also a to do list when students first log in showing what is due soon. Along the timeline, students first submit a document. Many document types are allowed, and they can be automatically converted to PDF by the system to ensure students can read the documents easily. The document submission deadline is actually a soft deadline: students can submit a document late with a per-day penalty that teachers can choose to override if there is a valid reason, such as illness or individualized education plan. Documents may be submitted up until one day before the final reviewing deadline.

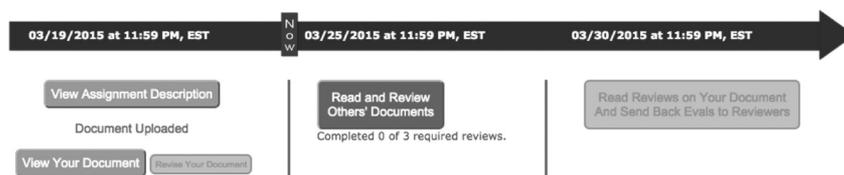


FIGURE 14.1 The Student Assignment Time Line View. Blue buttons show optional actions/past steps, green buttons indicated what action needs to be done next, and greyed out buttons show future steps.

Second, students review between three and six peer documents. Each document is labeled only by the author's chosen pseudonym (e.g., cubsfan15 or ForestGump); real identities of authors and reviewers are always kept hidden. Teachers determine the minimum number of reviews to complete for that assignment (four or five are recommended). Students may complete extra reviews for bonus points. For a given review, students must type comments in reaction to specific comment prompts. And for each comment prompt, there are one or more rating rubrics. For example, Figure 14.2 shows a comment prompt relating to the quality of the evidence provided and a corresponding rating prompt. The rating prompts are always on a seven-point scale, but they can be customized for how many of the points on the scale have concrete anchors. The students can scroll through the submitted document while completing the review, or they can download the document for printing. Each assignment's reviewing form can have as many or as few evaluation dimensions as the teacher would like. Typically, teachers include four to eight different reviewing dimensions in a given assignment. Assignments will vary in how much time it takes to do the reviewing task, but for typical high school writing tasks, students generally self-report spending 30 minutes to an hour on all the reviewing work.

Finally, students as authors examine all the feedback they have received and rate each reviewer's comments for helpfulness on a 1–5 scale (called back-evaluations), along with a brief explanation for the rating. Students generally value comments that contain at least some critiques (i.e., are not all praise), are polite in their criticism, and include suggestions for improvement. This back-evaluation step can be done quite quickly (e.g., typically in five to ten minutes).

3. Evidence for claims

Provide feedback on how well the author supported his specific and accurate references to the text.

Comment 1: (*Required)

Comment 2:

Evidence for claims
How strong is the textual evidence for each claim about Adams's rhetorical strategies?

XXXXX SELECT RATING XXXXX

7 - Every claim has accurate evidence for all important aspects of the claim. Most evidence is conveyed through direct quotes.

6 - 6

5 - Every claim has evidence, but some of the evidence is not accurate or not complete. Some evidence is conveyed through direct quotes.

4 - 4

3 - Several claims are missing evidence, or most of the evidence is not accurate. Little or no evidence is conveyed through direct quotes.

2 - 2

1 - No evidence is provided for any of the claims.

Rhetorical Analysis Essay: Jane Adams

At the time Jane Adams wrote this letter in 1780, she was an influential woman that truly portrayed a voice of this time by advising her son about the obstacles of life and persevering through them. To successfully influence her son to take her advice and show she supports him, Jane Adams utilizes the rhetorical analysis strategies of using several shifts to display strengths, allusions to establish credibility, and analogy to develop a perspective.

Jane Adams utilizes several contrasting shifts to compare her advice with stories that portray strengths and alliances that she hopes her son will be inspired to follow. Adams follows a consecutive shift from her own advice to something else such as narration. The first shift Adams demonstrates is a shift from giving advice to narration. Adams presents a real life encounter with an author (16), because she finds the advice the author gave her could help her son understand that his "improvement should bear some proportion to [his] advantages" (23-24). The next shift

FIGURE 14.2 Reviewing form showing text boxes for constructive feedback, one corresponding rating rubric, and a scrollable viewer for reading the document to be reviewed.

Students as reviewers can then see what back-evaluations their received for each review (see Figure 14.3). They can also see what other reviewers said and the ways in which the authors did or did not appreciate these other reviews. In this way, students can see models of good reviews and receive feedback on their own reviewing skills.

Students as reviewers can also see how accurately they rated each document on each rating dimension. Figure 14.4 shows what students see. For each rating dimension, the papers the student reviewed (five in this case), are ordered from lowest to highest according to what the other reviewers thoughts; thus the dark lines always have an increasing slope. In this example, the student’s own ratings and the others align well for explaining evidence and organization dimensions, but there is lower agreement for evidence for claims. Sometimes students have opposing views. In the case of high conflict, students can click to see what the other reviewers said to understand why they had a different evaluation of a given document.

The Teacher Tasks: Assignment Setup & Performance Monitoring

The teacher interface in SWoRD has all the basic learning management system functions one would expect. Teachers can create classes, and then inside classes, create assignments, provide feedback to students (strategically as they see fit), and monitor student and assignment performance.

There are some general parameters than can be adjusted, such as number of reviewers, grading policies (e.g., late penalties), but all the parameters have default values so relatively little needs to be done to get started except for entering the assignment description and choosing/creating evaluation dimensions.

The commenting prompts and rating rubrics can be created from scratch, selected from a library of shared prompts and rubrics, or selected from the teacher’s history of past prompts and rubrics. In addition, whole prior assignments can be copied, as can whole past courses; in fact, teachers can share a whole course’s content with another teacher simply by letting them know the enrollment key for their course.

	Reviewer #1	Reviewer #2	GreatGatsby	Reviewer #4
Explaining evidence	Goes in detail with explaining the evidence and ties it into the thesis statement.	The analysis of the evidence was sufficient enough for each section but there were varying strengths of their explanations. For example the paragraph about "personal accounts" is much weaker than the paragraph about "imagery". Develop the analysis of the personal accounts paragraph to create a more balanced essay.	The explanation was spot on such as in the connecting rhetorical questions to the evidence. It states how it connects to Lou's argument, and then to the authors interpretation.	You interworked the quotes to where they flowed and weren't abrupt in the sentence. I dont really have any complaints until the 4th paragraph, where I dont see any quotes. I like how you explain it, its just lacking textual evidence. Other than that you picked good evidence for everything else.
Backevaluation(3):	Mostly praise and no critiques.	Backevaluation(5): Very detailed and helpful!	Backevaluation(3): Thanks! How could I have furthered my textual evidence?	Backevaluation(4): How could the 4th paragraph be better? What quotes could I use.

FIGURE 14.3 Authors rating helpfulness of received comments on a 1 (unhelpful) to 5 (helpful) scale, along with brief explanations. Here the reviewer named Great Gatsby can see the helpfulness ratings they received along with the comments and helpfulness ratings of the other reviewers for this same document.

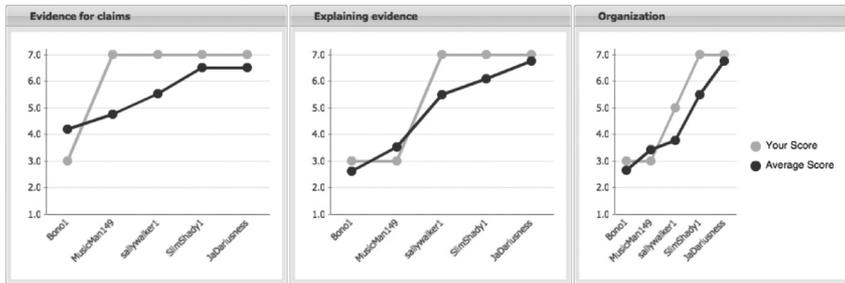


FIGURE 14.4 Comparison of a given student's ratings to the mean ratings from the other peer reviews on those same documents.

SWoRD makes it possible for students to do all the work on their own, receiving automatic email reminders as deadlines loom or if a paper or review is now late. But it is also useful for teachers to have an overview of the class as well as simple views of which students are struggling. Figure 14.5 shows the view of a particular assignment, providing information about overall submission rates and then listing particular students who missed particular deadlines or documents that overall have not received enough ratings or have high conflict across reviewers.

Of course, teachers can also see overall student performance on each of the rubric dimensions in terms of mean ratings. More interesting, however, is reliability performance of student ratings on each dimension: how well do students agree on a given dimension. Figure 14.6 presents an example from an English course. In this case, students had relatively low agreement on one of the evaluation rubrics, relating to describing the argument in an essay they analyzed. Two of the rating dimensions produced very high levels of inter-reviewer agreement, and the remaining dimensions were adequate, but still had areas that could be improved further. With this information, teachers can iteratively improve either their assignments or the in-class guidance they provide on the assignments. A recent analysis of all the high school writing assignments implemented in SWoRD found that a majority of the rubrics had good inter-rater agreement, suggesting that teachers quickly improve their assignments with this kind of feedback.

To understand why this dimension had lower agreement, a teacher can sort a table view of all of the submitted documents by level of agreement on that specific dimension, and from there jump to the reviews for the documents with the highest level of reviewer disagreement on that dimension (see Figure 14.7). For example, perhaps some students are not seeing certain problems at all. Or perhaps some students have a different view of how problematic certain issues are. The student comments provide teachers with insights into what the problems are, which can then lead to useful in-class discussions of the issues.

33/34 students submitted the document. Due: 02/26/2015
131/133 ratings completed. Due: 03/03/2015+1
131/133 comments completed. Due: 03/03/2015+1
121/131 backreviews completed. Due: 03/06/2015

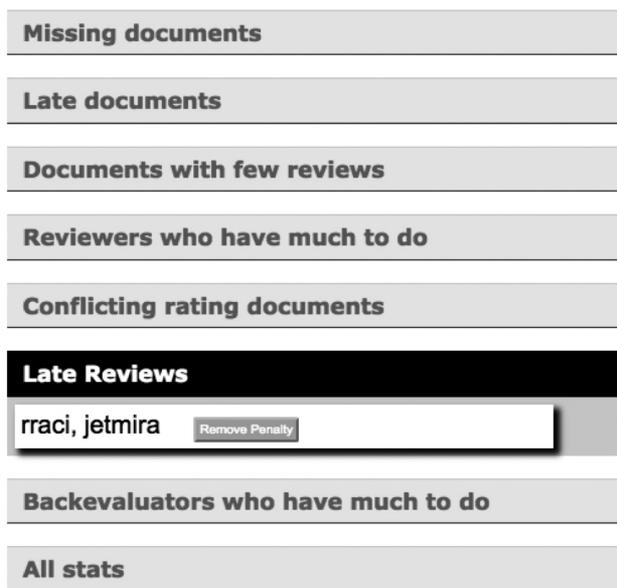


FIGURE 14.5 Teacher display of available student activity information on a given assignment. Clicking on a bar reveals which students have that issue (e.g., submitted a document late, or completed a late review). Late penalties for valid reasons can be waived with a click.

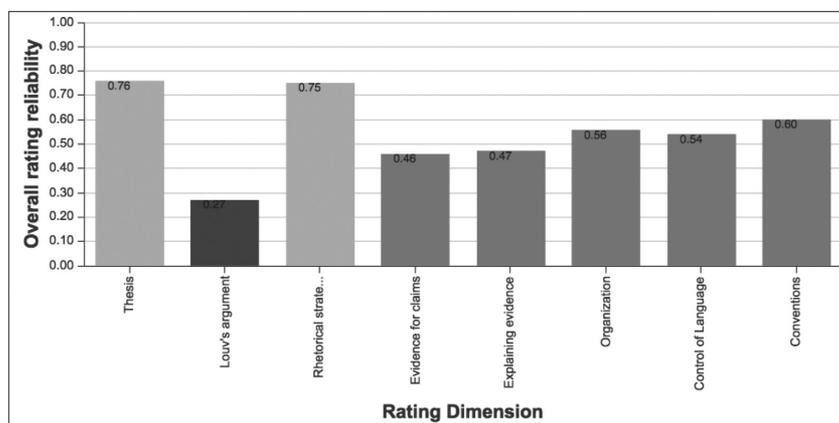


FIGURE 14.6 Teacher display of inter-rater agreement (also called rating reliability) of each rating rubric on a given assignment.

<p>Louv's argument</p>	<p>Within the first paragraph, Louv's argument is barely stated. Only the author's argument and opinion is stated. To fix the problem, use quotes from the text that relate to Louv's opinion on the separation between people and nature.</p> <p>Backevaluation(4): I should have more directly addressed Louv's argument. The reviewer's wording is a little strange but I understand the message.</p>	<p>The author doesn't ever truly state how Louv feels about the subject. They say that he believes that there is a growing separation between people and nature, yet never says how he feels. For example in the final paragraph they state that his essay is bringing "to light the thievery of real world appreciation that is being committed by the new technologies that we ourselves are creating." Here they are saying that Louv is showing this separation but yet again, not how he feels about it. The simple solution would be to include how he feels on the topic.</p> <p>Backevaluation(5): I agree that I didn't specifically state how Louv feels about the subject nor did I elaborate on his subject very much.</p>	<p>Argument is accurately and properly stated and explained. The separation between people and nature that Louv proves, is portrayed well in the essay.</p> <p>Backevaluation(3): It makes me feel good but doesn't really help me. Especially since other reviewers pointed out how I scarcely explained Louv's argument.</p>	<p>The opening to your Introductory paragraph was also very well written but I feel like it could have included more sentence explaining Louv's negative feelings about the separation and nature and his argument.</p> <p>Backevaluation(4): I agree I didn't outline Louv's argument very well in my introduction.</p>
-------------------------------	---	---	---	---

FIGURE 14.7 Reviewer comments in a case of high reviewer disagreement (along with author back-evaluations).

Research on Student use of SWoRD

Rating Accuracy

In general, peer ratings on sensible rubrics are similar to those produced by teachers, with there sometimes being a small bias in peer ratings being higher than teacher ratings (Falchikov & Goldfinch, 2000). The mean rating from four or five peers can be remarkably accurate when students are given a sensible rubric and incentives to take the reviewing task seriously. Across disciplines and student levels, the mean peer ratings from SWoRD were more similar to a teacher rating than any two teacher ratings would be to each other (correlations of .4 to .5; Cho, Schunn, & Wilson, 2006b). In a recent study (Schunn, Godley, & DiMartino, in press), we found that the mean of peer ratings in advanced placement (AP) classes using SWoRD were close to expert AP grader ratings that would meet the grading standards the College Board requires of AP expert graders. Even middle school students' peer ratings can be remarkably aligned to teacher ratings (e.g., correlations of greater than .9, or quadratic weighted Kappas of .85; Sadler & Good, 2006). These findings may seem surprising in that students clearly are generally not able to produce excellent work for the same tasks in which they are accurate raters; however, the high variability in quality across students, when combined with clear rating rubrics, makes assessing quality relatively straightforward. Further, students have received some instruction on what they should be doing (e.g., in the commenting and rubric prompts, as well in the classroom instruction leading up to the writing assignments). So, they often understand what should be done long before they have practiced enough to produce fluent and consistent outcomes in their own work.

Peers as an Important Audience

A concern that some teachers have is that students will not take an assignment seriously if they think it will only go to peers who might have lower standards. However, students are generally quite concerned with looking bad in front of peers. There is also the problem that teachers are usually an odd audience

for writing: students are asked to tell teachers things that it is clear the teacher already knows, which is a basic violation of norms of communication (Grice, 1975). Peers therefore might be a better audience than teachers for writing. A recent study with SWoRD found that first drafts of physics lab reports submitted to peers were actually stronger than first drafts submitted to teaching assistants (Cohen's $d = 0.5$), according to blind evaluations by experts (Patchan, Schunn, & Clark, 2011). A similar benefit from writing to peers rather than to a teacher was found in a study of writing at the middle school level (Cohen & Riel, 1989) suggesting it is a very general effect.

Comment Helpfulness

Although teachers generally look upon student peer feedback as inferior to teacher feedback, in fact, students commonly see peer feedback as just as helpful. This acceptance of peer feedback was found in student's helpfulness ratings of received feedback in a study in which students did not know whether feedback came from a peer or the teacher (Cho & Schunn, 2007). Similar results are shown in student surveys and interviews (Kaufman & Schunn, 2011; Topping, 2008). It is interesting to note, however, that students have significant concerns about the fairness of grades produced by peer review processes, even if they do not dispute the grades they received (Kaufman & Schunn, 2011; Sambell, McDowell, & Brown, 1997). More importantly, when looking at what students do with feedback they receive, peer comments can be just as helpful as teacher comments. While the teacher comments are more accurate, they are often not expressed in a way that students understand. For the more complex aspects of a task, we have often seen students make documents worse in response to teacher feedback, because of such misunderstandings (Cho & Schunn, 2007). Further, receiving comments from multiple peers in SWoRD can be more persuasive than feedback from just one individual, producing much larger improvements in the document from multiple peers than from one instructor (Cohen's $d = 1.2$; Cho & Schunn, 2007). When only one person makes a comment, a student can dismiss the comment as erroneous or the opinion of just one potentially biased person; when many peers make the same comment, even stubborn students acknowledge there is a problem. However, students still ignore many useful peer comments, just as they ignore many useful teacher comments (Nelson & Schunn, 2009).

Learning from Reviewing

Although it is important that the obtained ratings and comments are accurate and useful, a large benefit of implementing peer feedback is actually obtained from the act of providing constructive comments to peers. For example, one study found that students who provided feedback to others without themselves doing

any writing actually improved their own writing more than students who wrote and received feedback (Lundstrom & Baker, 2009).

Part of the learning benefit of reviewing peer's work is that students see models of what could be done as well as examples of errors that can be made (Cho & MacArthur, 2011). Sometimes this learning from seeing errors is about noticing the importance of the error (Kaufman & Schunn, 2011). For example, it is one thing to be told that dry writing is bad for communication; it is another thing to actually see how difficult it is to read dry writing.

But the benefits of reviewing are not just about seeing useful models. There is also the act of articulating the problem clearly and describing possible revisions that address the problem. In one study using SWoRD, we examined the value added of the commenting task in peer review. Some students were only asked to rate peer documents according to rubrics, whereas other students had to rate and comment. Subsequent submissions were significantly better in the rate and comment condition than the rate-only condition (Cohen's $d = 0.9$; Wooley, Was, Schunn, & Dalton, 2008).

It is for these reasons that SWoRD specifically rewards students for constructive reviewing through the helpfulness grades and that it allows bonus reviewing. Reviewing itself can be a strong learning opportunity, and having constructive comments is likely to be a critical aspect of maximizing that learning opportunity.

The Overall Benefits for Students

After using SWoRD, student and teacher surveys have regularly shown that students feel better prepared for writing-based exams, in part because they have had extra opportunities to practice, but more importantly because they have had the chance to think carefully about what is being expected of them (Godley et al., in press).

The prior sections showed the benefits of each aspect of the SWoRD process. How do they add up overall, and how broadly can SWoRD be used? There have not been careful studies of cumulative long-term impact, but one case study may be of interest. There is one high school that was an early and pervasive adopter of SWoRD, and it has obtained excellent exam results with over 95 percent of the students obtaining Advanced or Proficient scores on the 12th grade state writing test. This school involves many students coming from a background that is generally not associated with high exam performance (e.g., mean performance for other schools with 60 percent of students eligible for free or reduced lunch is below 70 percent Advanced or Proficient on this state writing tests). Yet, they were able to obtain writing performance that is more typically associated with the top performing suburban schools. These results were obtained from only two teachers using SWoRD across several years of instruction; their students wrote almost every day, and used peer review throughout the year. As a result of these excellent outcomes, the whole high school is moving toward regular use of

SWoRD, even extending it to math classes. Of course, one cannot make too much of a single case study and it is likely that these teachers implemented a number of helpful writing pedagogies. Further, this case provides no insight into how much peer review is required for obtaining good outcomes; but it does show that peer review can be used to greatly increase the amount of writing with feedback that students are asked to do. Additional research will be needed to establish how much and what kinds of peer review work are needed to produce writing gains in different kinds of subjects and in various educational contexts.

How to Access SWoRD, Now Named Peerceptiv

The first decade of SWoRD was paid for by research grant funding, and teachers throughout the world could use SWoRD for free. After a major rebuild of the system in 2009, which added many requested features and improvements to the interface, the number of SWoRD users was growing rapidly. The growth rate required the purchase of more web-servers and technical support, which were not supported by grant funding. Further grant funding is not a sustainable model for broad use of new technologies, because there can be gaps in funding for some years regardless of user demand. Since university researchers are neither skilled nor incentivized to distribute learning technologies broadly, and seeing the need to offer SWoRD on a very wide and sustainable basis of use, the University of Pittsburgh licensed SWoRD to a small company. The company renamed the tool Peerceptiv and implemented a number of additional improvements to the student and teacher interface; the user interface images shown earlier are all from the current Peerceptiv interface as of Fall 2015. There is still a version of SWoRD being used at the University of Pittsburgh for small research studies and testing new features to support student and teacher learning.

Peerceptiv is available via the web: <https://go.peerceptiv.com>. There is no local software installation. Any internet browser except for old versions of Internet Explorer will work. So, students and teachers can access the technology from home and school, as long as they have access to internet-enabled desktops or laptops.

Teachers and students create accounts from the main login page by click on the New Account button. Peerceptiv normally requires a valid email account, which is then used as the login ID and the destination of reminder emails for deadlines. (However, teachers can also set up a given class so that students only use teacher-provided usernames and passwords; in this case students will not get reminder emails.) At account creation, students select a pseudonym. The system checks to make sure they do not use their actual names. Because teachers can see these pseudonyms, no extra check is required that students do not select inappropriate names.

Teachers can create as many courses as they like within the system, either from scratch or copying prior courses they have created. For any schools using Blackboard, Brightside/D2L, Schoology, or Canvas, Peerceptiv will also directly integrate with those learning management systems (LMSs). Instructions on how

to set up that integration are available at the Peerceptiv website (see Help:LTI Configuration on the www.peerceptiv.com website). Through this mechanism, students in the class are automatically added into Peerceptiv as well, and grades are automatically passed back to the LMS.

Otherwise, teachers have two different ways of getting students enrolled in a Peerceptiv class. When a course is created, Peerceptiv generates a simple, easy to remember course key (e.g., able7). Students select “join a course” and enter the key once to then have that course associated with their account. Alternatively, teachers can also upload a file of student names and email addresses, and Peerceptiv will add them to the course, create accounts for students who do not yet have one, and email the students to have them finish the account creation process (e.g., picking a pseudonym). It may be important to check that the school or district system administrators will allow access to Peerceptiv.com and allow student email accounts to receive email from info@peerceptiv.com.

Access to Peerceptiv can happen through four different models. First students (or parents) can pay a course fee. Second, teachers can pay a per-student fee for the course (unlimited use throughout the year). Third, a department or school can purchase a pack of student seats that can be shared with multiple teachers. Fourth, a school can purchase a site license that allows an unlimited number of students on courses for all teachers using the school email address for their account.

Ways of Integrating Peerceptiv into a Classroom

Teachers vary greatly in the intensity of use of Peerceptiv. Some teachers use Peerceptiv two or three times a semester, whereas others have a new peer review assignment every other week. A few teachers have even done a new assignment every week, but the workload there is quite high, and students can get lost in all of the required steps when the pace is that rapid.

For students new to the technology, teachers typically have students create accounts in class, submit the first document in class, and begin doing the reviewing in class. Later students are able to do most of the work from home or other locations they use for computer access (e.g., the library). When new rubrics are introduced, it can be helpful to discuss as a whole class how a couple of example solutions match the rubric. Alternatively, teachers can use the performance diagnostics and only discuss the problem areas when they occur.

A common method is for teachers to have first drafts assessed only by students, and then second drafts be assessed by teachers, perhaps with only minimal comments. When teachers first use Peerceptiv, they are often skeptical of using the peer ratings for grading purposes. But after a few rounds of comparing their grades to the peer-provided grades, teachers quickly come to see that the mean ratings are also accurate for their classes.

Some early in-class discussions about what makes for useful comments can make the very first round of using Peerceptiv successful. Alternatively, through the

back-evaluation process, students usually come to see what kinds of comments are seen as helpful and what common problems produce unhelpful comments.

Conclusion

SWoRD/Peerceptiv is a web-based tool for implementing effective peer review assignments in a broad range of disciplines, seeking to implement writing-based assignments. A challenge for most teachers that normally prevents them from allocating sufficient assignments to produce improvements in student outcomes is that they do not have the time to provide feedback to students. Peer review, by contrast, does not burden the teacher in this way, especially when implemented using automatic distribution methods as found in SWoRD. Traditional implementation of peer review produces short superficial feedback, but using the structures and accountability mechanisms found in SWoRD, students generally produce accurate ratings and helpful comments. Further, a number of studies show that students learn from providing meaningful feedback to their peers.

There are still a number of open areas of work, however. While students feel that peer feedback is quite helpful (both providing and receiving), they are quite suspicious of grades produced by peer review, even though objective studies find these grades to be quite accurate. Additional research is required to find methods for overcoming these student concerns. Many students still ignore the peer feedback comments they receive, and thus student uptake of feedback remains a concern here just as it is a concern in their use of teacher feedback. Finally, very little is known about the cumulative impact of extensive use of such automated peer review as a pedagogical tool. Studies on this topic will need to be done in the coming years as this kind of peer review finally achieves sufficient use.

References

- Applebee, A. N., & Langer, J. A. (2011). A snapshot of writing instruction in middle schools and high schools. *The English Journal*, 100(6), 14–27.
- Biber, D., & Conrad, S. (2009). *Register, Genre, and Style*. Cambridge, MA: Cambridge University Press.
- Cho, K., & MacArthur, C. (2011). Learning by reviewing. *Journal of Educational Psychology*, 103(1), 73–84.
- Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education*, 48(3), 409–426. doi:10.1016/J.Compedu.2005.02.004.
- Cho, K., Schunn, C. D., & Charney, D. (2006a). Commenting on writing—Typology and perceived helpfulness of comments from novice peer reviewers and subject matter experts. *Written Communication*, 23(3), 260–294.
- Cho, K., Schunn, C. D., & Wilson, R. W. (2006b). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Journal of Educational Psychology*, 98(4), 891–901.

- Cohen, M., & Riel, M. (1989). The effect of distant audiences on students' writing. *American Educational Research Journal*, 26(2), 143–159.
- Dignath, C., Buettner, G., & Langfeldt, H. P. (2008). How can primary school students learn self-regulated learning strategies most effectively? *Educational Research Review*, 3(2), 101–129. doi:10.1016/j.edurev.2008.02.003.
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70(3), 287–322.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics: Vol 3, Speech Acts* (pp. 43–58). New York: Academic.
- Hayes, J. R., Flower, L., Schriver, K., Stratman, J., & Carey, L. (1987). Cognitive processes in revision. In S. Rosenberg (Ed.), *Advances in Applied Psycholinguistics Vol. 2. Reading, Writing and Language Processing*. Cambridge, UK: Cambridge University Press.
- Katstra, J., Tollefson, N., & Gilbert, E. (1987). The effects of peer evaluation on attitude toward writing and writing fluency of ninth grade students. *Journal of Educational Research*, 80(3), 168–172.
- Kaufman, J., & Schunn, C. (2011). Students' perceptions about peer assessment for writing: Their origin and impact on revision work. *Instructional Science*, 39(3), 387–406. doi:10.1007/S11251-010-9133-6.
- Kellogg, R. T. (1994). *The Psychology of Writing*. Oxford, UK: Oxford University Press.
- Kellogg, R. T., & Whitford, A. P. (2009). Training advanced writing skills: The case for deliberate practice. *Educational Psychologist*, 44(4), 250–266.
- Kiuhara, S. A., Graham, S., & Hawken, L. S. (2009). Teaching writing to high school students: A national survey. *Journal of Educational Psychology*, 101(1), 136–160.
- Kliwer, W., Lepore, S. J., Farrell, A. D., Allison, K. W., Meyer, A. L., Sullivan, T. N., & Greene, A. Y. (2011). A school-based expressive writing intervention for at-risk urban adolescents' aggressive behavior and emotional liability. *Journal of Clinical Child and Adolescent Psychology*, 40(5), 693–705. doi:10.1080/15374416.2011.597092.
- Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing*, 18, 30–43.
- Miller, L. D. (1992). Teacher benefits from using impromptu writing prompts in algebra classes. *Journal for Research in Mathematics Education*, 23(4), 329–340. doi:10.2307/749309.
- National Assessment of Educational Progress. (2011). *Writing Framework for the 2011 National Assessment of Educational Progress*. Retrieved from <http://www.nagb.org/publications/frameworks/2011naep-writing-framework.doc>.
- National Commission on Writing. (2004). *Writing: A Ticket to Work . . . Or a Ticket Out: A Survey of Business Leaders*.
- National Commission on Writing in American Schools and Colleges. (2003). *The Neglected R: The Need for a Writing Revolution*. New York: College Board.
- Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science*, 37(4), 375–401.
- Palincsar, A., & Brown, A. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1, 117–175.
- Patchan, M. M., Schunn, C. D., & Clark, R. J. (2011). Writing in natural sciences: Understanding the effects of different types of reviewers on the writing process. *Journal of Writing Research*, 2(3), 365–393.
- Patchan, M. M., Hawk, B., Stevens, C. A., & Schunn, C. D. (2013). The effects of skill diversity on commenting and revisions. *Instructional Science*, 41(2), 381–405. doi:10.1007/S11251-012-9236-3.

- Rubin, R., & Carlan, V. G. (2005). Using writing to understand bilingual children's literacy development. *Reading Teacher, 58*(8), 728–739. doi:10.1598/Rt.58.8.3.
- Sadler, P., & Good, E. (2006). The impact of self-and peer-grading on student learning. *Educational Assessment, 11*(1), 1–31.
- Sambell, K., McDowell, L., & Brown, S. (1997). "But is it fair?": An exploratory study of student perceptions of the consequential validity of assessment. *Studies of Educational Evaluation, 23*(4), 349–371.
- Schunn, C. D., Godley, A. J., & DiMartino, S. (in press). The reliability and validity of peer review of writing in high school AP English classes. *Journal of Adolescent & Adult Literacy*.
- Siegler, R. S. (1996). *Emerging Minds: The Process of Change in Children's Thinking*. New York: Oxford University Press.
- Topping, K. J. (2008). Peer assessment. *Theory Into Practice, 48*(1), 20–27.
- Van Den Berg, I., Admiraal, W., & Pilot, A. (2006). Peer assessment in university teaching: evaluating seven course designs. *Assessment and Evaluation in Higher Education, 31*(1), 19–36.
- VanDeWeghe, R. (2004). Awesome, dude! Responding helpfully to peer writing. *English Journal, 94*(1), 95–99.
- Vygotsky, L. S. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.
- Wooley, R., Was, C., Schunn, C., & Dalton, D. (2008). *The Effects of Feedback Elaboration on the Giver of Feedback*. Paper presented at the 30th Annual Meeting of the Cognitive Science Society.

