Mind the gap! Students' use of exemplars and detailed rubrics as formative assessment

Article *in* Instructional Science · July 2014

DOI: 10.1007/s11251-013-9299-9

4 authors, including:

Anastasiya Lipnevich
CUNY Graduate Center
40 PUBLICATIONS **911** CITATIONS

SEE PROFILE

Leigh Shebanie McCallen
CUNY Graduate Center
4 PUBLICATIONS **23** CITATIONS

SEE PROFILE

Jeffrey K Smith
University of Otago
102 PUBLICATIONS **1,637** CITATIONS

SEE PROFILE

Instr Sci (2014) 42:539–559
DOI 10.1007/s11251-013-9299-9

Mind the gap! Students' use of exemplars and detailed rubrics as formative assessment

Anastasiya A. Lipnevich · Leigh N. McCallen · Katharine Pace Miles · Jeffrey K. Smith

Received: 25 February 2013 / Accepted: 11 November 2013 / Published online: 21 November 2013

Abstract The current study examined efficient modes for providing standardized feedback to improve performance on an assignment for a second year college class involving writing a brief research proposal. Two forms of standardized feedback (detailed rubric and proposal exemplars) were utilized is an experimental design with undergraduate students (N = 100) at three urban college campuses. Students completed a draft of a proposal as part of their course requirements and were then randomly assigned to receive a detailed rubric, proposal exemplars, or a rubric and proposal exemplars for use in revising their work. Analyses of students' writing from first draft to second draft indicated that all three conditions led to improvements in writing that were significant and strong in terms of effect size, with the stand-alone detailed rubric leading to the greatest improvement. Follow-up focus groups with students indicated that a stand-alone rubric potentially engages greater mindfulness on the part of the student. Practical implications are discussed.

Keywords Feedback · Formative assessment · Classroom assessment

The use of formative assessment to enhance student performance and achievement has undergone a renaissance in recent years, leading to a variety of studies examining aspects of the relationship between formative feedback and students' ability to profit academically from such feedback (Evans 2013; Hattie and Timperley 2007; Kingston and Nash, 2011; Shute 2008; Symonds 2004). Researchers generally agree that comments specific to an individual's work and unaccompanied by grades tend to be most conducive to

A. A. Lipnevich (✉)
Queens College and The Graduate Center, CUNY, New York, NY, USA
e-mail: a.lipnevich@gmail.com

L. N. McCallen · K. P. Miles
The Graduate Center, CUNY, New York, NY, USA

J. K. Smith
University of Otago, Dunedin, New Zealand

improvement (e.g., Lipnevich and Smith 2009a, b; Hattie and Timperley 2007). In an extensive meta-analysis of the research on assessment feedback in higher education, Evans (2013) points out the critical importance of how feedback is received by the learner, how it is used by learners, and the incredible demands that the provision of feedback can put on instructors. This last point motivated the research presented here: The time and resource demands involved with the provision of assessment feedback. We were primarily concerned with an issue that rarely appears in the formative assessment literature: Can teachers deliver good quality feedback in a manner that does not require inordinate amounts of time? The purpose of the research reported here was to examine the relative efficacy of two approaches to providing feedback that do not require intensive work by teachers. Specifically, we investigated the impact of providing non-individualized, standardized feedback that was the same for all students on a particular assignment. Through the design of the study, issues of the motivational response of students to feedback, and the provision of the opportunity to engage the gap between current and desired states of learning (Ramaprasad 1983; Sadler 1989) were also considered.

Alternatively known as assessment *for* learning and formative assessment, assessment feedback is a pedagogical framework designed to promote learning and engagement (Black and Wiliam 2009). Working from Ramaprasad's (1983) conceptualization developed for work in management theory, Sadler (1987) proposed that the core of formative assessment comprises the determination of the gap between the actual state of performance and the desired state of performance, as well as individuals' engagement in efforts to close the gap. That is, in order for assessment to facilitate learning, students need to understand the level and nature of their current performance, the desired state of proficiency, and the discrepancy between the actual and the desired state. Further, they need to be able to effectively process that information and work to reduce the difference, or borrowing from the London Underground, "mind the gap." By performance, we mean the work that the student has produced in an instructional or assessment setting. In the study presented here, it was a brief research proposal written by the students, but it could refer to any production of material that is representative of a student's current state of ability and effort.

Typically, information on the current state of student performance is provided by teachers based on their assessment of students' work (Ilgen and Davis 2000; Kluger and DeNisi 1996). This is often thought to be the heart of formative assessment. But this way of thinking is neglectful of the two other components in the process: a clear understanding of what the goal or target is, and an environment conducive to engaging the gap between current and desired status. We argue here that it may be the case that learners at a more advanced level—for example, college students—can effectively assess their own current state of performance if provided with detailed information on the desired state, and incentive to use this information to improve their work. If this is the case, then students will be able to determine the gap on their own, and work to improve their performance. They will grow not only in the content area of the work under consideration, but perhaps also in their ability to self-assess. These are the ideas we pursue in this research.

Feedback within the framework of formative assessment

The efficacy of feedback in the instructional process is generally well accepted. Reviews of the literature and meta-analytic work on feedback have generally drawn the same conclusion over the past quarter of a century: feedback works (Black and Wiliam 1998; Crooks 1988; Hattie and Timperley 2007; Kluger and DeNisi 1996). However, there are

subtleties about how feedback works that are sometimes lost in the more general finding of feedback's overall effectiveness. In fact, meta-analytic work suggests that feedback may negatively affect performance in up to one-third of cases (e.g., Kluger and DeNisi 1996; Bangert-Drowns et al. 1991). In their meta-analysis of highly rigorous research in formative assessment, Kingston and Nash (2011) found that formative assessment practices are more effective in language arts than in mathematics or science. Further, Kluger and DeNisi (1996) contended that when feedback was accompanied by praise or critical judgments, the effectiveness of the feedback decreased, and that feedback that showed participants how to reach correct solutions was more effective than simple dichotomous judgments of correct/incorrect outcomes. Similarly, Bangert-Drowns et al. (1991) found that although feedback was positively related to greater achievement in most settings, student performance did not improve if feedback messages failed to include information necessary for learners to evaluate where they are, where they are going, or did not provide useful strategies to get them there. This conclusion is consistent with that of a number of researchers in the field—feedback that encourages "mindfulness" is most likely to help students improve (Hattie and Timperley 2007; Underwood and Tregidgo 2006). That is, comments that prompt students to meaningfully and thoughtfully approach revisions tend to result in the highest gains in performance.

Highlighting a key aspect of effective formative feedback from their review, Hattie and Timperley (2007) suggest that written feedback must encourage active processing of information on the part of the learner. As noted by Lipnevich and Smith (2009a), if students do not successfully engage with the feedback that they receive, feedback will not enhance student learning. In an experimental study conducted in situ in a large introductory psychology class, Lipnevich and Smith (2009a, b) found that detailed written feedback from the instructor without grades or praise was the most effective form of formative assessment for improving performance on an essay exam. Students in this condition (as compared to those being informed that their feedback was computer-generated, and those receiving a preliminary grade and/or a statement of praise and encouragement) showed the most improvement from a preliminary draft to a final revision. Follow-up focus group discussions with students who participated in the experiment revealed that students saw grades as potential obstacles to improvement, particularly by those who believed they had received a grade from the instructor. Students considered praise pleasant but the least influential form of feedback, useful only for balancing the demotivating effect of grades. Taken together, these findings present strong evidence that providing students with individualized, descriptive feedback specific to their work, and providing a setting where working on revisions based on that information could lead to an improved grade, can result in significant improvement in writing performance.

Formative assessment in higher education: challenges and opportunities

Assessment in a higher education context comes with a number of particular challenges and opportunities. On the challenge side of the equation, class sizes are burgeoning with courses frequently enrolling over a hundred students (Bose and Rengel 2009; Nicol and Macfarlane-Dick 2007). Added to this is the fact that classes typically meet only once or twice a week, and usually for the duration of a semester rather than a full year. To implement formative assessment effectively, instructors must return feedback in a timely and individualized manner, a feat hard to accomplish when teaching a large number of students. Second, and not frequently addressed in the literature, there simply is not a tradition of formative assessment in higher education, either from instructors or students.

Although there has been some excellent work in this area recently (see, e.g., Jonsson 2013), and of course McKeachie's classic work on teaching in higher education (Svinicki and McKeachie 2012) contains excellent suggestions on formative assessment, formative feedback is not common compared to summative feedback, and when it is provided, it is often underutilized or not utilized at all (Brown and Glover 2006; Sinclair and Cleland 2007).

On the opportunity side, many college-level instructors acknowledge the benefits of formative assessment practice (Bailey and Garner 2010), even though they feel they do not have the time to actualize what they consider to be best practice. Also, at the college level, students are more mature, have a more refined set of learning strategies, and are better able to process feedback than are elementary or secondary students (see, e.g., Lipnevich and Smith 2009 as compared to research by Andrade and her colleagues detailed below). Thus it may be the case that college students are better prepared to process formative feedback if presented under optimal conditions. Nicol and Macfarlane-Dick (2007) lay out a series of recommendations for good practice in formative feedback at the tertiary level. Among other suggestions, they argue that feedback should make clear what good performance looks like, facilitate the process of self-assessment among students, and provide opportunities for students to work on the area where they need improvement. The importance of promoting students' self-assessment skills is echoed in the work of Riordan and Loacker (2009) and Carless et al. (2011).

Feedback to improve student writing

The provision and use of feedback to improve writing can be seen as corresponding to the "rewriting" stage in a linear model of writing, and the "revision" or reviewing phase in a cognitive process model (Flower and Hayes 1981). Revision and formative assessment are dually concerned with processes and outcomes implicated in how a student engages "the gap", as stated in Fitzgerald's (1987) guiding definition: "revision means making changes at any point in the writing process. It involves identifying discrepancies between intended and instantiated text, deciding what could or should be changed in the text…changes may or may not affect the meaning of the text, and they may be major or minor" (p. 484). Where feedback is considered a keystone transaction to implementing formative assessment to improve student writing, revising is also considered a fundamental component of the writing process within a cognitive paradigm, such that re-writing is equivalent to the process of writing itself (Allal and Chanquoy 2004).

Flower and Hayes (1981) delineate a cognitive theory of writing process in which planning, translating, and reviewing are hierarchically organized phases that are managed by a writer's monitor (or meta-cognition), and influenced both by the task environment and the writer's long-term memory. Planning entails generating ideas, organizing, and goal setting; translating involves putting ideas into written language; and reviewing involves the sub-processes of evaluating and revising. Hayes et al. (1987, as discussed in Hayes 2004) later elaborated a more complex model of the revision process, proposing a sequence involving the writer's task definition (the goals and scope of the revision), model of evaluation (reading of the text to comprehend, evaluate, or define problems, depending on the writers' goals), and subsequent strategy selection process to respond to these goals and demands. A shortcoming of this model, noted by Hayes (2004), is the focus on problem solving cued by a writer's error detection as opposed to the pursuit of opportunities for new connections or ideas, such that "problem detection becomes a necessary pre-condition for

revision". The author cites the need for researchers to investigate the role that discovery plays in revision processes, an approach other researchers have identified as characteristic of experienced adult writers' revising (Sommers 1980). Hayes also reviews extensive evidence to suggest that writers at all levels can be taught to understand and apply criteria of text quality to improve their own written work, highlighting overall the recursive and continual role that revision plays throughout the writing process.

Though the cognitive processes involved in revising written work naturally implicate the role of feedback (whether generated by self, peer, or teacher), differentiating among these sources and the nature of the feedback generated is not a central question in studies seeking to conceptualize process models of writing and revising. In her review of the literature on revision in writing, Fitzgerald (1987) discusses some evidence contrasting the effects of teacher or peer feedback on revision processes versus the effect of teacher directions to revise (i.e. telling students to revise their work). Findings of studies reviewed tended to suggest that feedback can enhance revision, and that both peer feedback and teacher feedback is beneficial for improving writing quality, where as simply cueing students to revise their work has mixed or minimal value. The author also points to evidence showing that, for high school age students and more skilled writers, revising one's written work improves the quality of composition and contributes to writing achievement, yet little emphasis is placed on revising in American public schools. In our view, feedback plays a central role in how teachers engage students in effective cognitive and self-regulatory processes required to improve the quality of their written work.

The research presented here is not directly concerned with the improvement of student writing per se, but it does involve providing feedback to students on a written assignment in a psychology class. Educators play a crucial role in providing effective feedback to improve student writing, independent of the subject matter of the course. Teachers generally value the practice of giving feedback (Hyland and Hyland 2001) and actively use feedback in the process of teaching writing to students (Matsumura et al. 2002). The research on feedback on writing clearly indicates that the quality of feedback messages influences the extent of students' writing improvement (Reid et al. 2011; Wiliam et al. 2004; Ruiz-Primo and Furtak 2007). Kingston and Nash (2011) noted that the quality of feedback and the way it is used matters greatly and that the implementation of feedback is often "left to the discretion of the teachers implementing formative assessment" (Kingston and Nash 2011, p. 34). Thus, we are left with a dilemma: carefully constructed feedback messages on students' written work can lead to enhanced performance, but providing high quality feedback responses is time-consuming and may be impractical for teachers in many situations. Our goal in this research was to try to find a feasible approach to solving this dilemma.

Rubrics and exemplars in the framework of formative assessment

Andrade (2000, 2005) describes the use of instructional rubrics as a means to present students with the information on goals and aims of an assignment, thus taking care of one of the key components of formative assessment—that of knowing where you want to be (i.e., Black and Wiliam 1998; Sadler 1998). Andrade and her colleagues (Andrade and Boulay 2003; Andrade 2005) also looked at the possibility of using rubrics as a tool for fostering self-assessment by encouraging students to compare their work to a set of clearly specified criteria. Knowing clearly what good performance looks like and fostering self-assessment capabilities are two of the three goals of good formative assessment that were

postulated by Nicol and Macfarlane (2007). The Andrade team conducted a series of studies looking at the efficacy of providing rubrics to students on writing tasks (Andrade 2000, 2005; Andrade and Boulay 2003; Goodrich Andrade 2001). This research was primarily conducted at the middle school level and typically used a rubric at the beginning of an instructional sequence, finding rather modest results overall. Further, Reddy and Andrade (2010) reviewed the literature on the use of rubrics in higher education, again with somewhat mixed results. We feel that rubrics may be far more effective when used with college students who tend to be more meta-cognitively sophisticated than their middle school counterparts. We also think it may be more effective to use the rubric as a feedback mechanism after an initial draft has been turned in based on a more rudimentary explication of the assignment.

The use of exemplars has also been cited as a valuable method to aid students' understanding of marking criteria and subject standards (Sadler 1987; Orsmond et al. 2002). Exemplars (for instance, an "A" or a "B" grade essays) can be very effective indicators of where a student should be, and studies have demonstrated improvements in students' work when this tool is utilized. For example, Orsmond et al. (2002) report a study on self- and peer-assessment with students working on assessing a biology assignment using a set of exemplars. The study revealed that exemplars were effective in promoting higher quality of performance. Exemplars helped students to better understand standards of the assignment and, especially in the case of peer-assessment, formed a focus for meaningful formative feedback.

Further, Foster and Marasco (2007) suggested that combining rubrics and exemplars might result in superior performance as compared to using either of these tools separately. Presenting students with both rubrics and exemplars may allow them not only to know the assessment criteria for a writing task, but also to know what a finished piece of writing looks like at the different levels. In formative assessment parlance, it should help students to know exactly where they want to be when working on their assignment. To our knowledge, there are no studies that examined potential benefits of using rubrics or exemplars as a form of feedback. Our study intended to do just that—we investigated the effectiveness of rubrics and exemplars, alone and in combination, in promoting improvement on a writing task.

Aims of the present study

In an effort to bring efficiency into the equation for formative assessment while sacrificing as little of the effectiveness as possible, we have developed the following argument. In order for feedback to be effective, the literature suggests that three conditions must be realized (i.e., Black and Wiliam 1998). First, the student must have a sense of his or her current level of performance. Second, the student must have a strong sense of what the desired end state is so that the gap between current level and end state can be appreciated. Third, the student must actively engage in working toward the desired end state. The problem with meeting the first condition, providing extensive, individualized feedback to students, is that it is often simply too time-consuming for the instructor. So the question is, can the conditions necessary for effective feedback be realized while reducing the time investment of the instructor?

We propose a rather bold approach toward a solution. *Instead* of providing extensive individualized feedback on where a student currently is, teachers could provide extensive feedback on the desired level of performance without any feedback on where the student is,

and provide an incentive to the student to actively engage the material. Thus, the teacher provides the latter two conditions necessary for feedback to be effective, and lets the student engage in self-assessment to meet the first condition. To do that, we are proposing giving students an extensive rubric on how their work will be evaluated, and/or providing exemplars of poor, fair, and outstanding performance on the task. The students can then use this material to examine their own work and determine how it can be improved. This self-assessment can be "substituted" for extensive feedback on the current state of the student's work. As an incentive, we propose allowing students to revise their work and that the revised draft will be graded, not students' initial effort.

We realize that at first blush, this strains the definition of "providing feedback" or formative assessment, almost to the breaking point. However, let us consider what is being proposed here. Formative assessment, as it is typically realized, focuses on current status. Teachers mark papers with regard to mistakes, what needs improvement, what is currently good, etc. Usually there is not a lot of information on what outstanding performance might look like, or what the teacher is specifically looking for (this is not always the case, but often is so). Students often have a good idea of where they are (if they process the teacher's information), but not what a desired level of performance is. Also, usually there is no incentive to improve on drafts; what has been marked is often the end of the assessment process. In the model we investigate here, we emphasize the latter two conditions of good feedback: a clear sense of the desired state and an incentive to engage with the material. As to the first condition, we are leaving that up to the student. That is, we are relying on the student to be able to assess his or her own work against a detailed rubric and/or exemplars, judge the gap between current and desired state, determine what needs to done, and execute the changes on the work. Doing so successfully will result in a higher mark.

To summarize, we examined possible ways to provide effective feedback that would improve undergraduate student writing performance *and* be efficient for educators to deliver. To this end, we investigated student responses to detailed, non-individualized feedback in the form of an instructional rubric and research proposal exemplars that were given as tools for students to use to improve their final draft and grade on the assignment. Thus, we posed two research questions:

1. What are the effects of non-individualized feedback on students' performance?
2. What type of non-individualized feedback is more effective in promoting improvement?

Method

We addressed our research questions by employing an experimental design occurring within the context of an actual college course. Students were given the opportunity to revise an assignment that was part of their course requirements by using feedback they received following the first draft of their work. We randomly assigned students to one of three feedback conditions: (1) *Rubric*, in which students received a detailed description of how their work will be graded, broken down by different levels of performance; (2) *Exemplars*, in which students received three examples of the assignment, ranging from Weak to Average to Excellent; or (3) *Rubric and Exemplars*, in which students received both the rubric and the three exemplars. Upon receiving feedback, students were encouraged to use the materials to revise and resubmit their work. Their mark on the assignment was based on their revised work. This allowed us to study how important aspects of feedback influenced participants'

subsequent behavior in their efforts to improve their work. The primary dependent measure was students' score on the revised proposal, and the covariate was the score on the first draft of the proposal. We would like to note that we had originally intended to have a group that received neither rubrics nor exemplars, but could not receive permission to do so from the university IRB board where the experiment was conducted. This is a trade-off between ensuring the ecological validity of an in situ design and the ability to put participants in a control group that gets no feedback. The question arises, "What would happen if students received no feedback, but did have the opportunity to re-work their assignment?" Lipnevich and Smith (2009) were able to effectuate such a situation (in a more complexly designed study) and found that students receiving no feedback but an opportunity to rework an essay showed no improvement in their scores (Hedges $g = 0.012$, ns). Students with detailed feedback improved strongly ($g = 1.23$, $p < 0.001$). In this study, our primary goal was to look at the efficacy and efficiency of providing exemplars and/or rubrics, and decided to forego a comparison with a control group in order to have the experiment take place within a real instructional setting.

Participants

Participants in the experiment were second year students enrolled in child development courses at three college campuses at a large Northeastern urban university. The sample size for the experiment was 100 students, with 20 students attending campus 1, 61 attending campus 2, and 19 attending campus 3. Study participants ranged in age from 19 to 46, with a mean age of 22.5 years (SD = 5.3). Seventeen percent of the participants were male and 83 % were female. Of the participants, 35.4 % identified themselves as White, 24 % identified themselves as Asian/Pacific Islander, 6.3 % identified themselves as African American, and 34.4 % identified themselves as Hispanic. The mean self-reported GPA was 3.1 (SD = 0.50).

Procedures

As part of the course requirements, students were asked to write a 2–3 page research proposal demonstrating their basic understanding of methods and approaches used to conduct research in the field of child development. Prior to the assignment, three instructors delivered a lecture on research methods using the same set of presentation slides. Students were required to participate in the assignment as it was part of the course, but participation in the study was voluntary. All students agreed to participate. Because this task was a course requirement, the ecological validity of the study was enhanced. A copy of the assignment sheet was given out to all of the students after the presentation, and each of the three instructors reviewed the assignment with their classes. Students were asked to write about an issue, topic, or phenomenon in child development of their choice. The assignment page gave several relevant suggestions, as well as a list of required components regarding content, organization, and style guidelines. Information provided in the assignment guidelines included a list of the criteria delineated in the instructional rubric for the assignment (see below), but without the detailed description of the three performance levels. Students were told that they needed to electronically submit a first draft of their writing on a particular date, and then the instructor would email them materials to support the revision process. Students were also given a specific date to submit their second draft. Duration of time for writing the first and second draft was equivalent across classes. Students' score on the final proposal accounted for 10 % of their overall grade in the course.

After all first drafts were received, instructors used block randomization to ensure that an equal number of students in each class were assigned to the three feedback conditions. Class rosters were divided into three sets of names, and individuals in each set were assigned either the number one, two, or three based on the order provided from a random number generator. The three numbers represented the three feedback conditions: (1) Rubric; (2) Exemplars; (3) Rubric and Exemplars. Instructors individually emailed each student with their assigned form of feedback as an attachment. The format of the email was the same across classes for each condition. Students were asked to use the attached materials as a means to edit and revise their writing. Students were required to submit the second draft to their instructor via email 5 days after receiving feedback.

Two of the three experimenters graded both first and second drafts of the assignment using the instructional rubric (see Table 2). Students were not shown their draft scores. The final scores for first and second draft used in the analyses are averages of the two scores provided by the two raters. The agreement rate between the two raters was high (ICC = 0.94). After turning in the second drafts, students were provided will all forms of feedback (grading rubric and proposal exemplars) and were encouraged to revise their work again. A student's last draft submitted was used to count towards their final grade in the course, but the second draft was used as the dependent variable in the study. The maximum number of points for each draft was 30.

Rubric condition

Students in this condition received a rubric broken down into 10 criteria (see Appendix 1): (1) Description of Research Topic, (2) Study Design, (3) Study Materials, (4) Participants, (5) Procedures, (6) Implications of Potential Findings for Parents and Families, (7) Implications of Potential Findings for Someone Working with Children, (8) Format, (9) Writing/Grammar, (10) Overall Style. Each criterion was broken into three levels (corresponding to three point values) of performance: (1) Below Expectations, (2) Meets Expectations, (3) Exceeds Expectations. Key differences among the three levels of performance were presented in bold for clarity.

Exemplars condition

Students in this condition were given a set of three examples of student work (see Appendices 2, 3, 4). The three writing samples were generated based on the criteria and levels of the instructional rubric. Thus, the three writing samples were delineated as a Weak, Average, and Excellent research proposal.

Rubric and exemplars condition

Students in this condition received both a copy of the instructional rubric and the set of three exemplars.

Student feedback

After all students had been debriefed on the nature of the research, the instructors asked them to provide either oral or written feedback about the experience. Students were asked to participate in group discussions and were given an option to write down their thoughts.

Students from each experimental condition were asked to reflect on the effectiveness of the specific form of feedback they had received and explain why it did and did not help them during revisions. Students were also asked to describe how exactly they interacted with the feedback. Seven students chose to provide written comments, and responses of all participants who participated in an open group discussion were recorded by instructors.

Results

Analyses of differences in final scores

Analysis of covariance (ANCOVA) with the type of feedback (3 levels) as a factor and the grade for the draft of the proposal (before revisions) as a covariate, was used to examine differences in the second draft on the assignment. We initially conducted an analysis that included campus as an independent variable. We found no differences or interactions with campus (alpha = 0.05), so the campus variable was dropped from all subsequent analyses. Results of the test of equality of variance (Lipnevich and Smith 2009) yielded a non-significant result (F (2, 97) = 1.56, $p = 0.21$), indicating that the homogeneity of variances assumption was met; the Shapiro–Wilk test was also non-significant ($p = 0.466$) indicating that the assumption of normality was met as well. A preliminary analysis of variance was conducted on the first draft scores to ensure that the groups were not different at the start of the study. The ANOVA results were non-significant (F (2, 97) = 0.528, $p = 0.592$). ANCOVA revealed significant differences among the three conditions (F (2, 96) = 4.0, $p < 0.01$, $\eta^2 = 0.08$). Post-hoc analyses (Tukey's LSD) revealed that students in the *Rubric* condition ($M = 24.6$, SD = 4.1) did significantly better ($p < 0.05$) on their second assignment than their counterparts in the *Exemplars* ($M = 22.9$, SD = 3.9) and *Rubric and Exemplars* ($M = 22.3$, SD = 3.5) conditions ($p < 0.05$). There were no differences in performance between students in the *Exemplar* and *Rubric and Exemplars* conditions ($p = 0.09$). It is important to note that all three groups showed significant improvement from draft one to draft two. Effect sizes (pre- to post- for each group) ranged from 1.04 for the *Exemplars* and *Rubric and Exemplars* conditions to 1.54 for the *Rubric* only condition. See Table 1 for means and standard deviations. The results for the first and second drafts by condition are presented in Fig. 1.

Analyses of student feedback sessions

Students in the *Exemplars* condition received Weak, Average, and Excellent examples of a research proposal. Interestingly, students in this condition unanimously stated that they did not use Weak or Average examples when engaging in their revisions. A student noted: "I only looked at the strongest essay. Why waste time on weak ones? I want my proposal to be strong, so I modeled it after strong." Students felt that exemplars provided great examples of what their work should be like and felt that it was relatively easy to revise their proposals based on it: "It was really easy to make my proposal look like the one you [the instructor] provided." Participants remarked that having examples of great work is the easiest and most efficient way to ensure best outcomes in the revision process.

Similarly, students who received both rubric and exemplars stated that the exemplar of the Excellent work was what helped them most. A student voiced an opinion supported by all others, saying that if there is a choice between a rubric and an exemplar, he would

Table 1 Means, standard deviations, and effect sizes for the three experimental conditions

	N	Draft 1		Draft 2		Cohen's *d*
		M	SD	*M*	SD	
Rubric	36	18.9	3.3	24.6	4.1	1.54
Exemplars	32	19.0	3.6	22.9	3.9	1.04
Rubric and Exemplars	32	18.7	3.4	22.3	3.5	1.04
Total	100	18.8	3.4	23.3	3.95	1.22

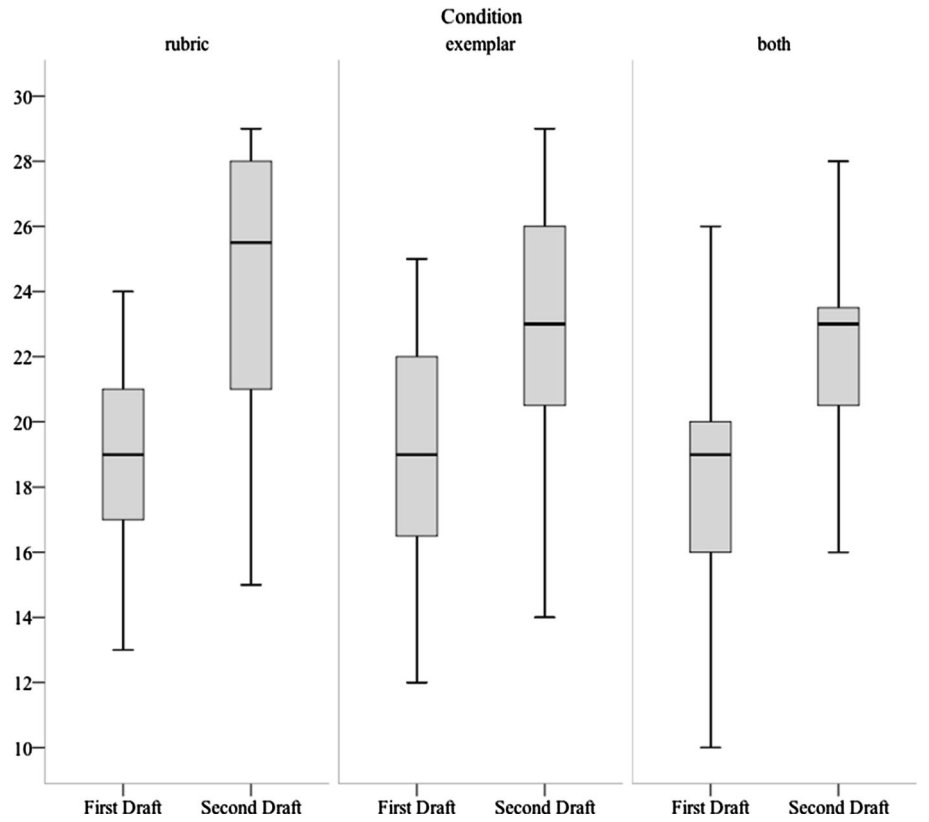

Fig. 1 *Box* and *whisker* plots of scores from students' first drafts and second drafts in the three experimental conditions

always go for the latter: "Exemplars are much more effective. They show exactly what your essay should be like. The rubric is very detailed, but exemplars are much more clear in what the professor expects." One of the students in the Rubric and Exemplar combined condition voiced a slightly different opinion: "I liked having the rubric. It told you exactly how the essay would be graded. Exemplars showed you what it should look like, so having both was very helpful." Some students in this group nodded in agreement, but restated the

higher perceived value of the Excellent exemplar, as compared to the rubric. Similarly to the participants from the *Exemplars* condition, students reported merely glancing at the Weak and Average example without interacting with them. "What's the point? You want your essay to be strong, not average or weak!" said one student. Most students who received exemplars echoed this view.

Students in the *Rubric* condition felt that this form of feedback was highly effective in helping them revise their work. A student proposed: "The rubric was very detailed, so I could go line by line, point by point and make sure I hit everything the way I was expected to." The participants reported that having a rubric made them break their proposals into separate components and work on making sure they appropriately addressed ever point. Students in the rubric group thought that having an example of the best proposal would have helped them to make revisions, but they did feel that rubric was greatly helpful.

Overall, the three groups agreed that having the opportunity to write and then revise their proposal was highly valuable. Students noted that due to demands of their lives, they often do not spend enough time writing and editing, and being encouraged to write and then re-engage with their work helped them to improve. Students appreciated the feedback and opined that having the Excellent exemplar was the form of feedback that was most effective and conducive to improvement. Students who were only presented with the rubric very much appreciated this form of feedback, but wished to have had an exemplar as well.

Discussion

The current study investigated the efficacy of providing detailed rubrics and written exemplars as a form of feedback to promote improvement in college students' written work. The results revealed that all three conditions led to improvement that was significant, and strong in terms of effect size. That is, giving students the opportunity to revise their written work, and providing them with information on how to improve, led to substantially enhanced performance.

These findings are generally consistent with the argument that feedback containing information on how an individual's work can be improved is highly effective. We note that the type of feedback used in our study carries no notion of how well the student had done up to that point. One might argue that because students were not informed about their current performance (i.e., grades, scores, qualitative evaluations, comments specific to their work), that they did not receive feedback at all. We feel, however, that presenting students with either a detailed rubric, or a set of exemplars, and encouraging them to evaluate their work against these materials created a very powerful formative assessment/feedback situation. The data from our study certainly suggest that this indeed was the case. It appears that students were able to assess their first efforts against either a rubric specifying grading criteria, or an exemplar of a good quality written essay, and make effective use of this information to improve their performance. We also note that the feedback provided did not involve social comparison or evaluation against a standard or norm. Thus, due to the fact that students' only sense of how well they had performed was based on their own assessment, it may well be the case that the negative affect associated with getting a grade (see, e.g., Lipnevich and Smith 2009a) was avoided in this approach. We would also like to emphasize that there were no negative statements contained in the feedback that students received. Whether this affected students' mood and motivation is speculative, as we did not measure affect in this study. It may be an avenue for future studies.

The effectiveness of our feedback may stem from the fact that neither rubric nor exemplar contained specific evaluation of student performance. Roos and Hamilton (2005) found that detailed neutral feedback is especially important for tasks that are loosely framed and do not have a clear "right" or "wrong" answer. The task employed in the current study fits the conceptualization of a loosely framed assessment of student performance (as compared to a multiple choice or short answer test). Not only did it require a strong command of the English language and good writing skills, it also required deep understanding and integration of numerous course-related concepts. The complex nature of this task helps to explain the role that feedback played in students' performance. Although students did not receive explicit information on where they currently stood with regard to the task, they did receive an explanation of what the expectation was (*Rubric*), or were presented examples of what good (or not so good) work really looked like (*Exemplars*), or both. As a result, it is possible that this kind of feedback prompted students to process information on a deeper level. After all, they had to decide whether they satisfied the requirements of the assignments, figure out what exactly they did wrong, and find ways to fix their work.

In terms of differences among experimental conditions, the *Rubric* condition produced the biggest net growth in performance, with students who received rubric alone generating a greater effect size by about 0.50, as compared to the other two conditions. Rubrics have been shown to be effective for communicating expectations to students (Andrade and Du 2005). Andrade (2005) notes that effective rubrics in writing clarify learning goals, guide educators' feedback on students' progress toward the goals, and allow students to judge their final writing product based on the degree to which they have met the learning goals. However, in her research, provision of rubrics as part of a feedback process was not as effective as the results seen here. We speculate that the essential difference in her findings and ours is that we were working with students who were older and perhaps more advanced with regard to general academic abilities (being college students). To our knowledge, no studies have previously examined rubrics as a form of feedback on a writing draft. Nicol and Macfarlane (2006) note difficulties in communicating directly to students what is expected of them in their writing, and argue that providing exemplars is a good way to address this problem. Indeed, we expected the exemplar condition to be the one that would return the most positive results, but the data indicate otherwise. Possibly, presenting rubrics only after students complete a draft of their assignment makes this tool even more effective as compared to offering it before students start writing (as used by Andrade and colleagues). This is something that future inquiries may want to explore. What is particularly interesting to us is that the *Rubric* only condition resulted in higher scores than the *Rubric and Exemplars* condition.

Focus group discussions that followed the experiment revealed that students focused almost solely on the best example when provided with exemplars, ignoring the poor and the average ones. Similarly, in the combined *Rubric and Exemplars* condition, students tended to only consider the best example, or focus on the exemplar more than the rubric. We believe that providing the rubric alone may have forced students to examine what they had done, and look to see how it met the requirements of the task, rather than trying to imitate the exemplar without checking their understanding of the task. This may represent what researchers have termed "mindfulness" (Bangert-Drowns et al. 1991). These authors argue that engendering mindfulness is the key feature in effective use of feedback. The results of this study demonstrated that the rubric may have called for a more sincere and

mindful engagement, which resulted in the student carrying out effective revision practices and thus improving their performance on the essay (although it should be kept in mind that students using exemplars also experienced substantial gains). Another possibility is that the rubric group showed the largest gains because the assignments were marked using the rubrics that the students were given. However, this would not explain why the rubric and exemplar performed similarly to the exemplar only group and not similarly to the rubric only group. The notion that the rubric engendered stronger engagement with the nature of the task seems a better explanation at this point.

Interestingly, although providing written feedback is a common form of writing instruction, according to Parr and Timperley (2010) feedback has not typically been a central theoretical concern in the literature on writing instruction. Thus, furthering our understanding of differential effects of feedback on students' learning is central to ensuring optimal outcomes of education (Lipnevich and Smith 2009a). This, combined with our finding that non-individualized feedback that is fairly easy for an instructor to deliver results in significant improvement, makes this study directly applicable to everyday teaching practices.

Limitations and future directions

Although the present study was strengthened by the in situ nature of the research, we acknowledge that what we have found here is that students made substantial improvements on the task "at hand." We do not know whether students receiving feedback on their written assignment would perform better in a subsequent task. One clear venue for future research is to examine how differential feedback influences subsequent learning in a course. It is, of course, difficult to conduct research that would vary the nature of the feedback that students receive on a randomized basis throughout an entire course, both for practical and ethical reasons. However, it may not be impossible. Further, although students emphasized the importance of revisions, a study could compare students receiving rubric and exemplars before they start working on their assignment to those who receive rubric and exemplars after they submit the first draft. It's not too fanciful to speculate that the second group would fare better, however, these differences should be tested. Additionally, future studies may examine qualitative changes in student writing depending on the feedback they receive. Finally, we need to consider the fact that there was no "control" group for this study. We do not know how well students would do if simply offered the opportunity to spend more time on their paper and submit a second draft.

Those limitations considered, we are greatly intrigued and encouraged by these results. They open up a realm of possibilities with regard to working with students to improve their skills, not only in the subject area under consideration, but their ability to self-assess, and to use that self-assessment productively in their educational endeavors.

Appendix 1

See Table 2.

Table 2 Rubric

		Below expectations (1 point)	Meets expectations (2 points)	Exceeds expectations (3 points)
1	Description of Research Topic	Research topic is not clearly stated or doesn't make sense. An explanation is either not given or a weak connection is made as to why this topic is important and relevant to the field of child development/language development	Research topic is clearly stated. An explanation is given as to why this topic is important and relevant to the field of child development/language development	Research topic is clearly stated and briefly explained. A thoughtful and convincing explanation is given as to why this topic is important and relevant to the field of child development/language development
2	Study Design	Student does not state the design (longitudinal, case study, observation, etc.) to be used in the study	Student states the design (longitudinal, case study, observation, etc.) to be used in the study; however, (1) the reason for using this type of design is not clearly explained; (2) a definition is not provided and/or (3) is not necessarily a good fit considering the research topic of interest	Student clearly states the design (longitudinal, case study, observation, etc.) to be used in the study. The reason for using this type of design is clearly explained, definitions provided, and is a good fit considering the research topic of interest
3	Study Materials	Student does not describe the materials (tests, checklists, questionnaires, etc.) to be used in the study	Student describes the materials (tests, checklists, questionnaires, etc.) to be used in the study; however the use of the materials is not clearly explained and/or the materials are not a good fit considering the research topic of interest	Student clearly describes the materials (tests, checklists, questionnaires, etc.) to be used in the study. The use of the materials is explained and the materials are a good fit considering the research topic of interest
4	Participants	Student does not state the participants (who and how many) to be included in the study or the majority of pertinent information regarding the gender, age, SES, ethical/racial, geographical, etc. demographics of the participants is missing	Student states the participants (who and how many) to be included in the study; however some pertinent information regarding the gender, age, SES, ethnic/racial, geographical, etc. demographics of the participants is missing	Student clearly states the participants (who and how many) to be included in the study. The projected gender, age, SES, race/ethnicity and other relevant characteristics are included
5	Procedures	Student does not state the procedures, or the student states only some of the procedures necessary to complete the study. Large gaps are evident in the procedure sequence leading to confusion of the reader	Student states the procedures, and explains the key steps. The reader has a general idea of the steps of the study	Student clearly states and explains the procedures necessary to conduct the study. The reader thoroughly understands the steps necessary to conduct the study

Table 2 continued

		Below expectations (1 point)	Meets expectations (2 points)	Exceeds expectations (3 points)
6	Implications of Potential Findings for Parents and Families	Implications are not clearly stated. Weak implications are mentioned that are irrelevant or trivial to parents and families	Relevant implications are stated in terms of how the findings might affect parents and families	Important implications of the findings are stated in terms of how the findings might affect parents and families. The reasoning between the importance of the findings for a parent/family is thoroughly explained
7	Implications of Potential Findings for Someone Working with Children	Implications are not clearly stated. Weak implications are mentioned that are irrelevant or trivial to people working with children (e.g., teacher, social worker, child advocate, etc.)	Relevant implications are stated in terms of how they might affect someone working with children (e.g., teacher, social worker, child advocate, etc.)	Important implications of the findings are stated in terms of how they might affect someone working with children (e.g., teacher, social worker, child advocate, etc.). The reasoning between the importance of the findings for someone working with children is thoroughly explained
8	Format	Format not followed Cover Page missing your name, course, title of the proposal, or date Paper is not typed or double spaced Margins are not 1 inch on all sides Font is not Times Roman 12 point type	Format Cover Page includes your name, course, title of the proposal, date Paper is typed, double spaced, Margins are 1 inch on all sides Font (Times Roman 12 point type)	Format APA format followed for Cover Page and includes your name, affiliation, title, and a running head Paper is typed, double spaced Margins are 1 inch on all sides Font (Times Roman 12 point type)
9	Writing/Grammar	Language has evidence of bias or stereotyping Writing is not clear and/or concise Use of incomplete sentences Improper use of punctuation Spelling errors are evident Lack of subject-verb agreement	Language has no evidence of bias or stereotyping Writing is clear Use of complete sentences Use of proper punctuation Use of proper spelling Consistent use of subject-verb agreement	Language has no evidence of bias or stereotyping Writing is clear Paragraphs are indented Use of complete sentences Use of proper punctuation Use of proper spelling Consistent use of subject-verb agreement
10	Overall Style	The proposal represents a series of disjointed ideas that are not organized into a cohesive narrative. The ideas does not flow from one point to another. Introduction and/or conclusion is missing	The narrative is rather smooth and cohesive, but introduction and conclusion is missing	The proposal is a cohesive, synthesized piece with a clear logic of presentation. There is a clear introduction and conclusion

Appendix 2

Excellent Proposal Exemplar

Running head: MAINSTREAMING VS SEGREGATING
Mainstreaming versus Segregating: Which Approach is Conducive to Development?
Jane Smith
Queens College

THIS IS A STRONG PROPOSAL, WITH STUDENT SCORING THE MAXIMUM NUMBER OF POINTS (30)

I would like to conduct a study to determine the effectiveness of special education programs as compared to inclusive classrooms for children with special needs. Specifically, I would like to focus on how children who have special needs can either fail or thrive in these classroom settings. I am hoping to be able to determine which setting will bring the greatest results for children with special needs.

In order to successfully and efficiently carry out this study, I will use several methodological approaches. First, I will use experimental design, in which I will randomly assign participants to two classrooms: inclusive and special education. Experimental studies allow us to conclude whether one of these instructional approaches causes improvement in students. I will administer a series of pretests, measuring motivation, self-efficacy, emotion, as well as subject matter knowledge. 6 month later I will retest my participants on all the measures to compare whether one of these groups did better.

Second, I will track my students longitudinally. I will administer the aforementioned assessments twice a year for a period of 3 years in order to see in which environment students fare better. The same students will participate. Longitudinal design will allow us to make conclusions about gradual changes occurring over time.

Third, I will use naturalistic observation in both special education classrooms and inclusion classrooms. I will use a checklist to record students' behaviors. I will record how special needs children are doing academically and socially in both settings; and analyze my findings to see which one works better for children with special needs. Observations will be conducted for a week every 6 months by the same experimenter. Although observations are rather subjective, they will provide us will rich data on students' in-class behaviors.

Participants will include 50 children with documented qualification for special services that are 7 years old, and will be studied until they turn 10 years old. I will try to select participants of various ethnic backgrounds, races, and SES. I will try to recruit 25 male and 25 female students.

The implications of studying this issue in terms of how it will affect a family is extremely important because of the rise of special needs children diagnoses. As of 2010, The Centers for Disease Control and Prevention lists the prevalence rate for autism, for example, as 1 in 110 children. If 1 in 110 children are diagnosed with autism, and there are so many other disabilities, the amount of families with special needs children that will need guidance is monumental. This study will enable parents of children with special needs to make informed decisions on which classroom environment will best suit their child and enable them to flourish, learn, and grow to the best of their abilities.

The implications of this issue in terms of how it will affect myself and professionals who will be working with children is vital as well. I believe one of the most important things in being an educator is to understand that every child, with special needs or not, is different and therefore learns differently. All educators need to understand that enabling

children to learn to the best of their abilities requires flexibility and personalization. My study will show the importance of how different children thrive in different environments, and show educators how to accommodate and understand every child that passes through their classroom doors to the best of their abilities. This study may have serious policy implications. Superintendents of large districts may decide to introduce more special needs classrooms or inclusive classrooms, depending on the findings.

In conclusion, my study on whether special needs children succeed better in inclusion or integrated classrooms is of extreme importance to both families, educators, and most importantly the children themselves. This study will enable teachers and other school professionals to understand whether children with special needs learn better in certain environments. It is of extreme importance, and could be life changing for the future education.

Appendix 3

Average Proposal Exemplar

Mainstreaming versus segregating: Which Approach is Conducive to Development?

Jane Smith
Queens College

THIS IS AN AVERAGE PROPOSAL, WITH STUDENT SCORING BETWEEN 18–22 POINTS

I would like to conduct a study to determine the effectiveness of special education programs as compared to inclusive classrooms for children with special needs. People send their children to all kinds of schools, so parents should know which one is the best. For children with special needs, being in a classroom that fits their needs can make a big difference for how well they do in school.

In order to carry out this study, I will use several methods. First, I will use experimental design, in which I will lace participants into two classrooms: inclusive and special education. I will measure their motivation, self-efficacy, emotion, as well as subject matter knowledge. 6 month later I will retest my participants on all the measure to compare whether one of these groups did better. Second, I conduct longitudinal design. I will measure motivation, self-efficacy, emotion, as well as subject matter knowledge twice a year for a period of 3 years in order to see in which environment students fare better. Third, I will use naturalistic observation in both special education classrooms and inclusion classrooms. I will use a checklist to record students' behaviors. I would like to study 50 children with special needs. The participants will be 7 years old, and will be studied until they turn 10 years old.

The implications of studying this issue in terms of how it will affect a family is extremely important because of the rise of special needs children diagnoses. As of 2010, The Centers for Disease Control and Prevention lists the prevalence rate for Autism, for example, as 1 in 110 children. If 1 in 110 children are diagnosed with Autism, and there are so many other disabilities, the amount of families with special needs children that will need guidance is monumental.

The implications of this issue in terms of how it will affect myself and professionals who will be working with children is vital as well. I believe one of the most important things in being an educator is to understand that every child, with special needs or not, is

different and therefore learns differently. All educators need to understand that in order to enable children to learn to the best of their abilities requires flexibility and personalization. My study will show the importance of how different children thrive in different environments, and show educators how to accommodate and understand every child that passes through their classroom doors to the best of their abilities.

In conclusion, my study on whether special needs children succeed better in inclusion or integrated classrooms is of extreme importance to both families, educators, and most importantly the children themselves. This study will enable teachers and other school professionals to understand whether children with special needs learn better in certain environments. It is of extreme importance, and could be life changing for the future education.

Appendix 4

Weak Proposal Exemplar

Mainstreaming versus segregating: Which Approach is Conducive to Development?
Jane Smith

THIS IS A WEAK PROPOSAL, WITH STUDENT SCORING LESS THAN 10 POINTS.

I would like to study how good special education programs are. Also, if they are better than mainstreaming. To study special needs is very important. Children who have special needs need to have attention from parents and teachers. It will be very important to know what kind of classroom is better to them.

In order to carry out this study, I am going to observe children in school and at home and see whether students in mainstreamed classes do better or do worse than children in special education classes. Observations are good because you see what actually happens in classes. It is very important to observe children because you can also notice a lot of different things. You can also observe in childrens home. I will also interview teachers and parents and ask them questions about who does better. I will study boys and girls who have different special needs. Maybe I will ask 10 childrens.

The importance of studying this issue is extremely important because of the rise of special needs children diagnoses. Parents will appreciate if they know what they do. Nowadays people do not whether special education classes or mainstreamed classes are good for your child. That's why the study is very important. For children it will also be important.

The importance of this issue in terms of how it will affect educators is also important. Because educators would like to know what to do with special needs children to help them learn a lot of things and grow.

References

Allal, L., & Chanquoy, L. (2004). Introduction. In L. Allal, L. Chanquoy, & P. Largy (Eds.), *Revision: Cognitive and instructional processes*. Norwell, MA: Kluwer Academic.

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership, 57*, 13–19.

Andrade, H. G. (2001). The effects of instructional rubrics on learning to write. *Current Issues in Education, 4*(4).

Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching, 53*, 27–30.

Andrade, H. G., & Boulay, B. A. (2003). Role of rubric-referenced self-assessment in learning to write. *The Journal of Educational Research, 97*, 21–30.

Andrade, H. G., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research, & Evaluation, 10*, 1–11.

Bailey, R., & Garner, M. (2010). Is the feedback in higher education assessment worth the paper it is written on? Teachers' reflections on their practices. *Teaching in Higher Education, 15*, 187–198.

Bangert-Drowns, R. L., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*, 213–238.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*, 7–68.

Black, P., & Wiliam, D. (2009). Developing a theory of formative assessment. *Educational Assessment, Evaluation, and Accountability, 21*, 5–31.

Bose, J., & Rengel, Z. (2009). A model formative assessment strategy to promote student-centered self-regulated learning in higher education. *US-China Education Review, 6*(12), 29–35.

Brown, E., & Glover, C. (2006). Evaluating written feedback. In C. Bryan & K. Clegg (Eds.), *Innovative assessment in higher education* (pp. 81–91). London: Routledge.

Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education, 36*(4), 395–407.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*, 438–481.

Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research, 83*(1), 70–120.

Fitzgerald, J. (1987). Research on revision in writing. *Review of Educational Research, 57*(4), 481–506.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*(4), 365–387.

Foster, G., & Marasco, T. L. (2007). *Exemplars: Your best resource to improve student writing*. Markham, ON: Pembroke.

Hattie, J., & Timperley, H. S. (2007). The power of feedback. *Review of Educational Research, 77*, 81–112.

Hayes, J. R. (2004). What triggers revision? In L. Allal, L. Chanquoy, & P. Largy (Eds.), *Revision: Cognitive and instructional processes*. Norwell, MA: Kluwer Academic.

Hyland, F., & Hyland, K. (2001). Sugaring the pill: Praise and criticism in written feedback. *Journal of Second Language Writing, 10*, 185–212.

Ilgen, D., & Davis, C. (2001). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology, 49*, 550–565.

Jonsson, A. (2013). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education, 14*(1), 63–76.

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice, 30*, 28–37.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: Historical review, meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*, 254–284.

Lipnevich, A. A., & Smith, J. K. (2009a). Effects of differential feedback on students' examination performance. *Journal of Experimental Psychology: Applied, 15*, 319–333.

Lipnevich, A. A., & Smith, J. K. (2009b). I really need feedback to learn: Students' perspectives on the effectiveness of differential feedback messages. *Educational Assessment, Evaluation, and Accountability, 21*, 347–367.

Matsumura, L. C., Patthey-Chavez, G. G., Valdés, R., & Garnier, H. (2002). Teacher feedback, writing assignment quality, and third-grade students' revision in lower-and higher-achieving urban schools. *The Elementary School Journal, 103*(1), 3–25.

Nicol, D. J., & Macfarlane-Dick, D. (2007). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218.

Orsmond, P., Merry, S., & Reiling, K. (2002). The use of exemplars and formative feedback when using student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education, 27*, 309–323.

Parr, J. M., & Timperley, H. S. (2010). Feedback to writing, assessment for teaching and learning and student progress. *Assessing Writing, 15*, 68–85.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science, 28*, 4–13.
Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education, 35*, 435–448.
Reid, J. L., Drake, S., & Beckett, D. (2011). Exploring teacher and administrator perceptions of assessment in the face of new assessment policies. *Canadian Perspectives: Education Coast to Coast to Coast*, *1*(1).
Riordan, T., & Loacker, G. (2009). Collaborative and systemic assessment of student learning: From principles to practice. In G. Joughin (Ed.), *Assessment, learning, and judgement in higher education*. Dordrecth: Springer.
Roos, B., & Hamilton, D. (2005). Formative assessment: A cybernetic viewpoint. *Assessment in Education: Principles, Policy & Practice, 12*, 7–20.
Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching, 44*, 57–84.
Sadler, D. R. (1987). Specifying and promulgating achievement standards. *Oxford Review of Education, 13*, 191–209.
Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education, 5*, 77–84.
Shute, V. J. (2008). Focus on formative assessment. *Review of Educational Research, 78*, 153–189.
Sinclair, H. K., & Cleland, J. A. (2007). Undergraduate medical students: Who seeks formative feedback? *Medical Eduction, 41*(6), 580–582.
Sommers, N. (1980). Revision strategies of student writers and experienced adult writers. *College Composition and Communication, 31*(4), 378–388.
Svinicki, M. D., & McKeachie, W. J. (2012). *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers* (Fourteenth ed.). Belmont, CA: Wadsworth.
Symonds, K. W. (2004). *After the test: Closing the achievement gaps with data*. Naperville, IL: Learning Point Associates.
Underwood, J. S., & Tregidgo, A. P. (2006). Improving student writing through effective feedback: Best practices and recommendations. *Journal of Teaching and Writing, 22*, 73–97.
Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education, 11*, 49–64.