Student Reflection to Improve Access to Standards-Based Grading Feedback

Heidi A. Diefes-Dux
School of Engineering Education
Purdue University
West Lafayette, IN, USA
hdiefes@purdue.edu

Laura M. Cruz Castro
School of Engineering Education
Purdue University
West Lafayette, IN, USA
lcruzcas@purdue.edu

***Abstract*— In this Research-to-Practice Full Paper, two different reflection prompts in the form of unstructured and structured reflections were implemented in an engineering course to guide students towards improved self-regulated learning behaviors, specifically accessing feedback on their assignments. The course setting for this study employed standards-based grading (SBG) which provides rich feedback on students' proficiency with the course learning objectives. Low student access to feedback, delivered through a learning management system (LMS), was seen as a considerable learning opportunity loss. Negative binomial regression models were used to investigate whether semester (2016 or 2017) and refection (presence or absence) had an impact on the (1) number of days students accessed the LMS gradebook or rubrics, (2) the number of times students accessed the gradebook to view grades, or (3) the number of times students accessed rubrics to view feedback. Semester, which relates to implementation changes in the SBG system, significantly increased students' number of days of access and number of gradebook and rubric accesses. Reflection, particularly structured reflection, significantly improved students' access to the rubrics. Pairing reflection with a well-developed SBG system has the potential to improve students' access to feedback.**

Keywords—self-regulated learning, standards-based grading, assessment

I. Introduction

Many studies have demonstrated that feedback can be powerful with regards to student achievement [1]. Feedback has been shown to be "more strongly and consistently related to achievement than any other teaching behaviour" [2, pp. 277-78]. Further, effective feedback can aid first-year students' transition to the university and may support student retention [3]. Ultimately, formative feedback is intended to improve student learning [4]. But "for assessment to be formative the feedback information has to be used" [5, p. 16]. Use of feedback requires, as a first step, reading feedback. Therein lies the problem.

An instructor who subscribes to the belief that "feedback is teaching" [6] or feedback is for learning [7] puts considerable time and thought into rubric design, grading, and the provision of (written) feedback. If students are minimally able to monitor their learning and fail to access their feedback, they miss a considerable number of opportunities to improve their learning. When a course employs standards-based grading (SBG), which provides rich feedback on students' abilities with regards to course learning objectives [8, 9], the learning opportunity loss is that much greater.

Instructors often witness graded papers go immediately from their hands to their students' backpacks with barely a glance at the grade, much less the feedback, possibly never to be seen again [10]. It is perhaps even less obvious what happens to feedback when students' work is graded via a learning management system (LMS). In one study, as few as 39% of students in a course that implement SBG via an LMS agreed with the statement that they had reviewed their feedback on early-course assignments soon after they were released [11].

The use of reflection has the potential to instill in students a habit of accessing their feedback as the reflective practice could nurture self-regulatory learning behaviors. The purpose of this study was to examine the effect of two different types of periodic reflections on students' access to feedback delivered via an LMS and compare their access to feedback to that of students who did not participate in reflection.

II. Theoretical Framework

This work is anchored in the theories of self-regulation and reflection in the context of academic settings. Self-regulated learning (SRL) is a process whereby students orient their thinking, behaviors, and actions towards attainment of their learning goals [12]. Self-regulated learning involves cycling through three phases - forethought, performance, and self-reflection [13]. Forethought entails goal setting and planning. Performance involves self-control (use of strategies selected during the forethought phase) and self-observation (active engagement in recording, experimenting, or monitoring learning processes). Self-reflection takes on two forms: self-judgment and self-reaction. Self-judgment (or self-evaluation) is concerned with comparisons of one's self-observed performance against some standard. It might also be concerned with attributing cause to one's errors or successes. Self-reactions may entail feelings of satisfaction, withdrawal from or avoidance of opportunities to learn (defensive reactions), or adjustments to one's method of learning (adaptive reactions). These phases can be taught through instruction and modeling and are aligned with qualities of lifelong learning [13].

Reflection is concerned with thinking about one's current knowledge or understanding of an issue or problem to fulfill a purpose or achieve an outcome [14]. As described in greater

978-1-5386-1174-6/18/$31.00 ©2018 IEEE

detail by Turns et al. [15,16], the theoretical underpinnings of reflection are derived from a large body of works that includes Dewey's making meaning of experiences [17], Kolb's experiential learning cycle [18], and Schön's reflection-in-action [19].

In the academic context, the outcome of reflection is typically targeted to some aspect of learning. Reflection in engineering education is gaining traction as part of a greater effort to move towards student-centered teaching [20]. The Consortium to Improve Reflection in Engineering Education, with representation from 12 U.S. institutions [16], recognizes that "Reflection and the promotion of reflective techniques are becoming more important in engineering education because of the expanding need for diverse, adaptive, broad-thinking, and nimble engineering experts who can respond to the ever-increasing challenges that society faces" (cpree.uw.edu). As such, reflection is being used in a wide variety of ways in engineering courses, including teaching of ethical reasoning [21], enhancing leadership competencies [22], and adjusting students' commonly-held misconceptions about heat and mass transfer [23].

III. Research Questions

In this study, two different forms of reflection prompts (unstructured and structured) were implemented in a first-year engineering course to guide students towards better use of course resources, including accessing feedback on their assignments. These periodic self-reflections engaged students in self-judgement and planning [13]. Of the many potential outcomes of reflection, the prompts used in this study were intended to focus students' reflective attention on their learning, knowledge, and understanding of the course content and some form of action [24]. The hypothesis is that reflection prompts with a high degree of structure leading to students' self-evaluation with regards to their proficiency with the course learning objectives will result in greater access to feedback as a learning strategy (self-control in the performance phase [13]). The research questions for this study were: How often do students access their grades and SBG feedback in an LMS? Does reflection and its form impact students' frequency of access to their grades and SBG feedback?

IV. Methods

A. Setting and Participants

This study was set in a large (N~1550) 16-week required first-year engineering course offered in Spring 2016 and Spring 2017, heretofore referred to as only 2016 and 2017. The course goals centered on developing students' abilities to: (1) apply basic (MATLAB) programming concepts to the solution of engineering problems, (2) represent and interpret data in multiple formats, (3) develop, select, modify, and justify mathematical models to solve an engineering problem, (4) function effectively as a member of a team, and (5) demonstrate habits of a professional engineer. The course was divided into sections of no more than 120 students. Each section met for 110 minutes twice a week. All sections followed the same curriculum with common lecture materials and resources, assignments, and exams. The one departure was extra credit; instructors used their allocated extra credit points in a number of different ways. Grading was done through an LMS, specifically Blackboard Learn™.

B. Reflection Prompts

Two sections were selected from each semester for inclusion in this study. One section in each semester did reflections on their learning (treatment) as their extra credit; the other section did something else that was unrelated to reflection (control). The instructor that did the reflections and the instructor that did not were the same in each semester. The primary objective of the survey type reflections [20] in each semester, was to engage students in actively thinking about their learning and to encourage them in devising a plan to improve their learning.

In the 2016 reflection section, a variety of unstructured reflections were used, with most being end-of-class minute papers [25-27] asking students to think about the most important thing they learned that day and to convey any questions they still had on the day's topic. Of the 21 reflections done across the semester, five were targeted at getting students to access their SBG feedback via the LMS. In weeks 3 and 4 when problem set (PS) 01 and PS02 feedback were released, the students were explicitly asked to "Take a look at the details of our evaluation of your PS0[X] work via the learning objective based rubric" and respond to this series of prompts: "What do you believe you have and have not learned? How well do you think our evaluation of your work represents what you have and have not learned? What specific actions do you plan to take based on our evaluation of your work?" A qualitative analysis was done of these responses [28]. In week 8, following the return of Exam 1, the students were asked "Considering our evaluation of your Exam 1 work via the Blackboard learning objective based rubric: How well do you think our evaluation of your work represents what you have and have not learned? What specific actions do you plan to take based on our evaluation of your work?" In week 11, in preparation for Exam 2, students were asked, "From the assessment of your work (Exam 1 & PSs), what (learning objectives) LOs are high priority for you to work on learning? How will you work on learning these specific LOs?" In week 14, following the return of Exam 2 and PS08, students were asked "From the assessment of your work (Exam 2 & PS08), what LOs are high priority for you to work on learning? What specifically do you need to work on with regards to these LOs?"

In 2017, more structured reflections occurred typically at the start of the first class each week. In the second class each week, unstructured reflections, typically minute-papers were done. The structured reflections were timed such that the students had both received feedback on a problem set or exam in the last few days and had just submitted the next problem set prior to the start of class. If feedback had been returned, the reflections asked students if they had accessed their feedback via the LMS and what actions they had taken or planned to take based on that feedback. These reflections also asked students to rate their abilities with each recent learning objective. For instance, students were asked in week 4 to rate their ability with learning objective 04.01 Create a script that adheres to programming standards. The rating options were: I can do this on my own without referring to resources; I can do this on my own if I refer

to some resources; I need more practice with this; I need someone to help me understand and do this; or I am not sure what this means. These reflections concluded by asking the students to propose a course of action to improve their learning.

Of the 25 structured reflections done across the 2017 semester, 10 included the full set of prompts asking students about their access to their feedback, abilities with recent learning objectives, and plan of action. These occurred weekly from weeks 3 to 9 and 11 to 13. An additional three reflections, occurred in weeks 2, 14, and 15. Two of these asked the students to rate their abilities with the learning objectives and to devise a plan of action. The third only asked students to rate their abilities with the learning objectives.

Overall, the depth of the reflections in both semesters was intentionally low, meaning the nature of the reflection intervention was, at a minimum, technical (focused on the examination of one's skills and competencies) and, at a maximum, descriptive (encouraging analysis of one's actions) [29]. These reflection interventions were not intended to stretch students towards higher level reflection behaviors such as considering alternatives or multiple viewpoints.

C. Grading

Table I lists the graded activities for the course. The grading scheme was such that 1000 points were distributed across problem sets, exams, project milestones, pre-class quizzes, teamwork, and miscellaneous assignments (e.g., technology agreement and end-of semester-survey). In 2016 and 2017, SBG rubrics were used to assess student work on problems sets and exams. The project milestones were graded with non-SBG criterion referenced rubrics in 2016 and SBG rubrics in 2017. For the SBG rubrics, each rubric item was a course learning objective (LO) that was rated in 2016 as fully achieved (100%), partially achieved (80%), underachieved (50%), or no evidence (0%) and in 2017 as proficient (100%), developing (80%), emerging (50%), insufficient evidence (0%), or no attempt (0%). Each time a learning objective was assessed it was allocated some number of points as shown in Table I. Pre-class quizzes were mapped to learning objectives but were graded based on the number of learning objective-based multiple-choice quiz questions correct. Teamwork skills were assessed using the CATME behaviorally anchored rating scale for self- and peer evaluation [30] and verified by instructor observations, though there were a few teamwork skills related assignments graded based on completion. Attendance issues resulted in point deductions depending on the issue as outlined in the syllabus. Miscellaneous activities were primarily graded based on completion with some attention to the quality of the submission on select assignments. An additional 20 points were allocated to instructors for extra credit. These were also graded on completion with some attention to the quality of the submission on select assignments.

All grading was completed via Blackboard Learn™ using the gradebook or a rubric. Students could access their overall grades on assignments via a single gradebook page (My Grades). Approximately 80 different grades were posted each semester (before the extra credit activities). If a rubric was used to assess their work and students wished to see their performance on each item, they had to intentionally access the rubric. Problem sets, exams, and project milestones had rubrics associated with them. In addition, written feedback was included in the rubric for problems sets and project milestones. In the case of exams, written feedback was provided on the exam papers outside of the LMS. Students had 20 assignments in 2016 and 23 assignments in 2017 assessed with a rubric. Starting in week 3, a student could expect in a given week that at least 1 assignment would be returned to them that used a rubric and at least four other grades would be posted to the Blackboard gradebook (e.g., pre-class quizzes, attendance).

In 2016 students had one additional reason for accessing rubrics through Blackboard on a weekly basis. Generic rubrics [31] were provided with each assignment. So students could view these rubrics while completing an assignment. In 2017, this was not an option; rather students were provided a list of learning objectives with descriptions of evidence of proficiency. This list was provided in a separate document made available with each assignment. An example of an LO and its evidence is shown in Table II.

TABLE I. COURSE ACTIVITY WITH BLACKBOARD GRADEBOOK ENTRIES

Graded Activity	**Spring 2016**					**Spring 2017**				
	No.	***Rubric Type***	***No. Items Assessed per Activity***	***Points Per Item***	***Total Points***	***No.***	***Rubric Type***	***No. Items Assessed per Activity***	***Points Per Item***	***Total Points***
Problem Sets	8	SBG[a]	4-10 LOs[a]	1-3	105	12	SBG	5-10 LOs	1-2	120
Exams	3 (I[a]) 3 (T[a])	SBG	19-33 LOs	2-10	300	3 (I) 3(T)	SBG	23-25 LOs	2-14	360
Project Milestones	6	CR[a]	5-14 criteria	1-20	375	5	SBG	7-23 LOs	1-10	280
Pre-Class Quizzes	17	None	1-6 LOs	1	50	19	None	1-8 LOs	1	70
Teamwork	10	None	NA	NA	150	10	None	NA	NA	150
Miscellaneous	2	None	NA	NA	20	2	None	NA	NA	20
Attendance	29	None	NA	NA	-X[a]	28	None	NA	NA	-X

[a] I = Individual, T = Team, SBG = Standards-Based Grading, CR = Criterion Referenced, LO= Learning objective, -X = points deducted

TABLE II. COURSE LEARNING OBJECTIVE AND EVIDENCE EXAMPLE (SPRING 2017)

LO 07.00: Create and evaluate x-y plots suitable for technical presentation
Sub-LO 07.05: Format plots for technical presentation
Evidence: • Correct syntax for title • Correct syntax for xlabel • Correct syntax for ylabel • A descriptive title that references the problem context, the independent (x) variable, and the dependent (y) variable • Clear x-axis label with units • Clear y-axis label with units • Gridlines • Color and marker/line style(s) that are as specified or distinctive (when multiple data sets) • Properly formatted legend, when multiple data sets and/or models • X-axis scales match each other (when using subplots to compare data)

D. Data Collection and Cleaning

The data for this study were students' learning analytics, specifically clickstream data recorded as students navigated Blackboard to the gradebook or a rubric. The gradebook showed a student their score on assignments. A rubric showed the student how they performed on each LO assessed on an assignment and their written feedback on each LO. Data was recorded for every student's access to the gradebook or a rubric for the duration of the course. A record of each access consisted of a student's ID, the student's access to a type of data (gradebook, rubric), the timestamp of the access to the nearest minute, and a unique session ID which is created each time a student logs on to Blackboard.

This study focused on the period of time in which students could have accessed the rubrics to read their feedback and used that feedback to guide their future work. Only that data logged between the time the first assignment was returned to students at the start of week 3 and the last assignment was submitted for grading in week 16 were used. Depending on the exact date the first assignment was released for students to view, the period of time included in the study was between 94 and 96 days.

E. Data Analysis

One challenge of the data was that a rubric access was actually an access to the rubric system. So it was unknown which rubric a student was specifically accessing. A second challenge of the data set was that the rubric system was typically accessed through the gradebook; an access to the rubric system appeared as two clicks, one to gradebook and one to the rubric, in quick succession. Pairs of gradebook-rubric clicks were treated as a single access to only the rubric system if the timestamps were less than or equal to one minute apart (the precision of the timestamps). A third challenge was to distinguish truly unique accesses to either data type when a student was coming in and out of the gradebook and rubrics multiple times in a given login session. A threshold of five minutes between access recordings in the same login session was set to make this distinction.

Three different negative binomial models were fit to the data to analyze the influence of the two independent variables (semester and reflection) on students' navigation to the gradebook and rubrics. The semesters were Spring 2016 and Spring 2017; reflections were either done or not done. The dependent variables for the three different analyses were (1) number of days in which at least one access was made to the gradebook or rubrics, (2) number of unique accesses to the gradebook, and (3) the number of unique accesses to the rubrics. Analyses were conducting using R [32].

As is appropriate for data that are counts, several models and linked functions were examined (e.g., Gaussian, Poisson). Negative binomial models were found to be the most appropriate because the data was found to feature over-dispersion [33], meaning the variance was greater than the mean, and the residual deviance, a measure of the goodness of fit, was much greater than the degrees of freedom. The Akaike's Information Criteria (AIC) [34] was used to identify the model type that brought the residual distribution closest to the degrees of freedom.

The interaction between semester and reflection was not statistically significant at a 95% confidence level. This means that semester did not influence the students' access behaviors in the reflection and non-reflection sections differently. As such, interaction was not included in any of the models.

Equation (1) represents the three models.

$$\widehat{\log(y)} = \widehat{\beta_0} + \widehat{\beta_1}(Reflection) + \widehat{\beta_2}(Semester), i = 1,2,3\ldots \quad (1)$$

where y is the number of days of access, the number of unique accesses to the gradebook, or the number of unique accesses to the rubric for each student (i). The dependent variable y is assumed to come from a distribution in which the variance of y is related to the mean as follows for a negative binomial distribution:

$$Var(y) = \mu + \alpha\mu^2 \quad (2)$$

Here α is the dispersion parameter for the model.

Reflection and semester are binary variables. Reflection was assigned 0 if the student was in a non-reflection section and 1 if a student was in a reflection section. Semester was assigned 0 when a student was in a 2016 section and 1 if a student was in a 2017 section.

Equation (3) is easier to use when interpreting the results as is shows that it is necessary to apply the exponential function to the model parameters. This equation was used to compute the reported average number of accesses under conditions where there were significant differences between sections.

$$\hat{y} = e^{\widehat{\beta_0} + \widehat{\beta_1}(Reflection) + \widehat{\beta_2}(Semester)} \quad (3)$$

The null hypotheses for this study were:

H_o: Reflection had no significant effect on students' (1) days of access, (2) number of accesses to the gradebook, or (3) number of accesses to rubrics.

H_o: Semester had no significant effect on students' (1) days of access, (2) number of accesses to the gradebook, or (3) number of accesses to rubrics.

V. Results

Table III describes the students in the study sections in terms of their participation in reflection, their course completion, their access to the gradebook or rubrics within the study dates of the semester, and their inclusion in the study. In 2016, there were eleven students that never accessed the gradebook or the rubrics. These were not poor performing students. In fact, they all had final grades of C or better; their median final grade was an A.

TABLE III. Students in Participanting Sections

Sem[a]	Reflect[a]	No. Completed Course	No. Accessed Gradebook or Rubric	No. Included in Study
2016	Yes	78	72	72
	No	115	110	108
2017	Yes	70	70	70
	No	113	113	110

[b.] Sem. = Semester; Reflect = Reflection Done

When using the negative binomial model, the estimators, which are generated using the Maximum Likelihood method (ML) [35], can be sensitive to the presence of outliers in the sample. Outliers can increase the probability of a Type 1 or 2 error. Five outliers were detectable using the Cook's distance. The Cook distance [36] for a generalized linear model was obtained from the library "base" in R [37]. The criteria used for the outlier detection was any student with a Cook's distance above $4/N$ where N was the number of students in a section that accessed the gradebook or rubrics within the study time period.

The five outliers were in the non-reflection sections. Each of these student accessed the gradebook more than 200 times during the study period. For these students, there were particular days in which the student was identified as having accessed the platform more than 50 times. This number of accesses was relatively resistant to changes in the 5-minute threshold for detecting unique accesses. The reasons for these high access counts is not known but could be related to a variety or things, for instance, study behaviors around an exam or a platform issue for that student based on the mode of access (website or app). The data for these five students were removed from the analysis.

Below, the basic descriptive statistics are summarized in visuals of the distributions of the data for the four possible combinations of the covariates (semester and reflection). A red dashed-and-dotted line shows the mean of the section, while a green dashed line shows the median. The standard deviation is also included.

The estimation results for the three negative binomial model are presented in tables following each figure showing the distribution. The model parameters and their characteristics (standard error, p-value) are shown in the tables.

A. Days of Access

Fig. 1 illustrates the distribution of the number of days students accessed either the gradebook or rubrics in 2016 and 2017 for the sections that did and did not do reflection. For both the reflection and non-reflection sections, the number of days of access increased from 2016 to 2017.

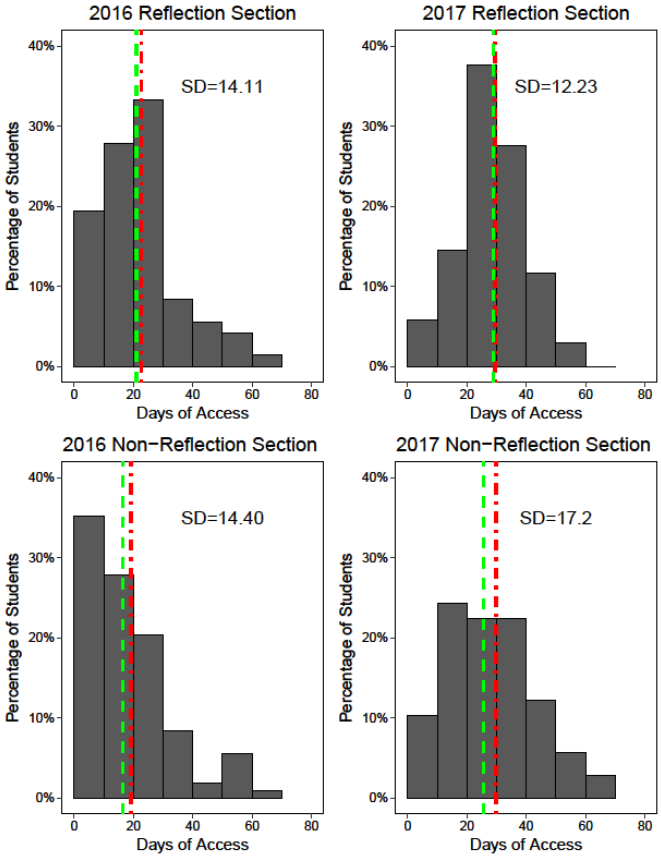

Fig. 1. Days of gradebook and rubric access.

This increase in the number of days of access is confirmed in the results of the negative binomial analysis as shown in Table IV. When comparing the number of days of access in 2016 to 2017, there was a significantly greater number of days of access in 2017, regardless of whether reflection was done or not. Students accessed the gradebook or rubrics on average 9 days more in 2017 than in 2016. For a given semester, the reflection versus non-reflection sections did not show any significant difference for the number of days of access.

TABLE IV. Negative Binomial Regression Results For Days of Access

Variable Description	Estimated Parameter (SE)	p-value
Intercept	2.989 (0.056)	< 0.0001
Reflection (Yes)	0.077 (0.070)	0.734
Semester (2017)	0.372 (0.069)	< 0.0001
Goodness of Fit	**Estimated Parameter (SE)**	
Dispersion α	2.753 (0.207)	
Degrees of freedom	357	
Residual Deviance	386.99	

B. *Gradebook Acccess*

Fig. 2 shows the number of times students accessed the gradebook in 2016 and 2017 for the section that did and did not do reflection. Like the number of days of access, the number of accesses to the gradebook increased from 2016 to 2017. There were a few students with more than 100 unique accesses to the gradebook. These students were only in the non-reflection sections.

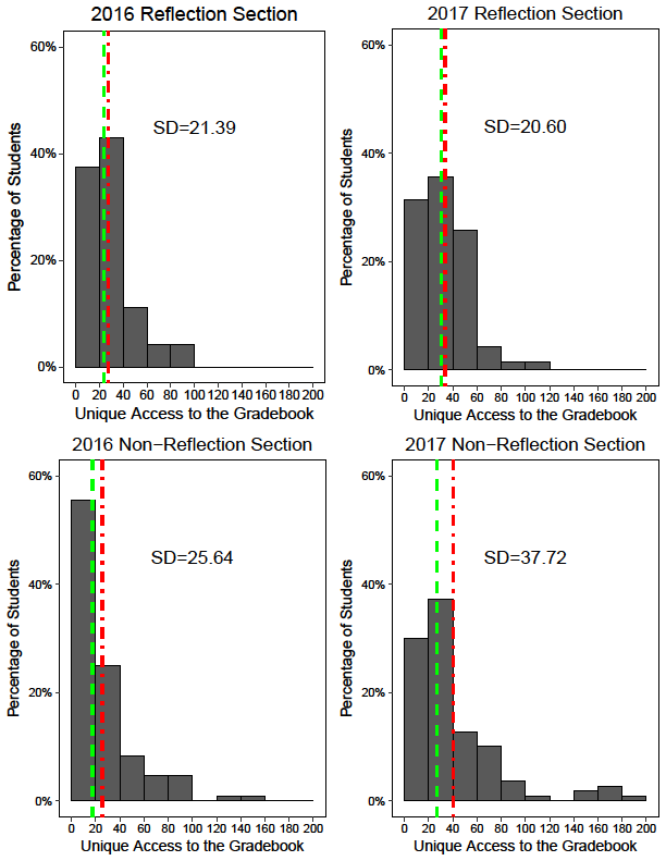

Fig. 2. Unique accesses to the gradebook

When comparing the number of gradebook accesses by semester, there was a significantly greater number of accesses to the gradebook in 2017, regardless of whether reflection was done or not (Table V). Students accessed the gradebook on average 11 times more in 2017 than in 2016. The reflection versus non-reflection sections did not show any statistically significant difference in the number of gradebook accesses.

C. *Rubric Access*

Fig. 3 shows the distribution of the number of times students accessed the rubrics in 2016 and 2017 for the sections that did and did not do reflection. A number of students never accessed the rubrics. There were 22 such students in the non-reflection section and 6 such students in the reflection section. In 2017, there were 2 such students in the non-reflection section and 1 student in the reflection section.

TABLE V. NEGATIVE BINOMIAL REGRESSION RESULTS FOR ACCESS TO THE GRADEBOOK

Variable Description	**Estimated Parameter (SE)**	**p-value**
Intercept	3.294 (0.070)	< 0.0001
Reflection (Yes)	-0.052 (0.087)	0.551
Semester (2017)	0.353(0.085)	< 0.0001
Goodness of Fit	**Estimated Parameter (SE)**	
Dispersion α	1.611 (0.12)	
Degrees of freedom	357	
Residual Deviance	393.69	

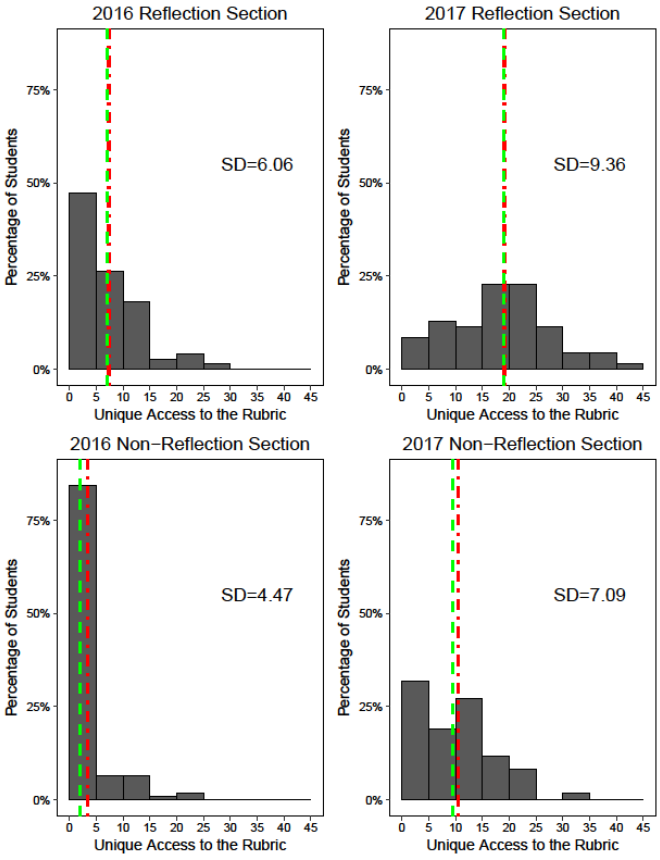

Fig. 3. Unique accesses to the rubrics

The result of the negative binomial regression analysis is shown in Table VI; both semester and year were significant. The 2016 reflection section students accessed the rubric on average 3 more times than the 2016 non-reflection sections students. The 2017 reflection section students accessed the rubric 9 more times than the 2017 non-reflection section students. In addition, students in the 2017 non-reflection section increased their rubric access. They accessed the rubrics 6 more times than their 2016 non-reflection counterparts. Finally, students in the 2017 reflection section accessed the rubrics 13 more times than those in the 2016 reflection section.

TABLE VI. NEGATIVE BINOMIAL REGRESSION RESULTS FOR ACCESS TO THE RUBRICS

Variable Description	**Estimated Parameter (SE)**	**p-value**
Intercept	1.259(0.077)	< 0.0001
Reflection (Yes)	0.684 (0.090)	< 0.0001
Semester (2017)	1.055 (0.088)	< 0.0001
Goodness of Fit	**Estimated Parameter (SE)**	
Dispersion α	1.784 (0.178)	
No. observations	357	
Residual Deviance	419.43	

DISCUSSION

What is readily apparent from the distributions of days of access to the gradebook or rubrics and the number of unique accesses to the gradebook and rubrics is that a large percentage of the students in 2016 in the non-reflection section were monitoring their performance rather infrequently. This group is significant because it represented the bulk of the students in this large FYE course in which the SBG system was nearly fully implemented and reflection was not being tried. For this section, 35% of the students accessed the gradebook or rubrics 10 or fewer days out of the 94 days (across 13 weeks) in which monitoring could have provided useful feedback for improved learning. Similarly, 55% of the students made 20 or fewer unique accesses the gradebook. Even more detrimental, 83% of the students accessed the rubrics 5 or fewer times when there was feedback available on 20 assignments across the semester that could have been useful to them.

To be fair, there were some issues with the SBG system in 2016 that may have led to poor student monitoring of their performance. There was a delay in the release of the grades on the first problem set that placed the availability of that feedback after the second problem set was due. It is not known whether this sort of delay was a persistent problem, but this first instance could have established a habit among the students of not checking for feedback prior to the next assignment due date. In addition, grader reliability was a significant issue, and the provision of written feedback when students' work was not deemed proficient was often missing or of insufficient quality to be helpful [38]. These issues could have reduced students' faith in the grading system, de-valuing it to the point of not motiving them to monitor their performance. Further, as a general course practice, feedback on any assignment was scheduled to be returned on the due date of the next assignment. The lack of time for students to look at feedback prior to the next assignment's due date may have led to their just ignoring feedback.

And yet, the modest number of reflections done in 2016 led to an improvement in the number of unique accesses to the rubrics. Only 47% of students accessed the rubrics 5 or fewer times. The first two reflections in particular attempted to engage students in reviewing their feedback. Early in the semester, this may have established a slightly better habit of reviewing ones feedback than in the non-reflection section. So even under less than ideal SBG implementation, one type of student performance monitoring behavior was positively shifted in the presence of a small amount of reflection. There is an alternative explanation. The reflection section's instructor was leading the design of the SBG systems implementation. This may have resulted in her making greater reference to feedback and reviewing ones performance in class and that alone might have contributed to her students slightly improved monitoring of their performance.

The significantly greater number of days of access to the gradebook or rubrics and the number of unique accesses to the gradebook in 2017 might be explained by a slight increase in the number of graded assignments from 78 in 2016 to 82 in 2017. More likely, the increase in performance monitoring, was the result of instructional changes around the implementation of the SBG system that produced a heightened awareness of the SBG grading system amongst the instructional team. First, as recommended by [38], a weekly grader training program was initiated to increase inter-rater reliability of the assessment of each LO and to improve written feedback. Second, SBG data results were provided to the instructors on a weekly basis indicating the performance of their students with regards to the LOs. Third, the SBG data were discussed each week in the instructional team meeting. The heightened instructional team awareness of the SBG system may have been transferred to the students, either implicitly or explicitly, in class.

In addition, in 2017, feedback on each assignment was made available to students at least two days before the next assignment was due, creating a greater window of time to learn from feedback before the next assignment was due. The grader training and the better enforced turn-around of feedback could have raised the value of the SBG feedback to the students and resulted in greater student monitoring of their performance.

The greater structure and frequency of the reflections in 2017 seemed to guide students to improved performance monitoring. Reflection in the 2017 section resulted in two very significant changes in students monitoring of their performance. First, the students accessed the rubrics on average nearly as many times as there were assignments with rubrics to view. The non-reflection section, on average, only accessed the rubrics 11 times when there were 20 assignments with rubrics to view. Second, the students in the reflection section did not have as much variance in the number of times they accessed the gradebook; their average accesses were similar to their median accesses. This means that the reflection section had fewer students with a high number of gradebook accesses. Perhaps, the more compulsive checking of the gradebook seen amongst some of the non-reflection section students was converted to more productive engagement with the rubrics in the reflection section.

A closer look at the 2016 and 2017 reflections using the features of reflection outlined by Turns et al. [16] can elucidate the reasons for the improvement in performance monitoring from 2016 to 2017 and highlight how the reflections can be further improved. The Turns et al. [16] reflection framework includes seven features: experience, feature, lens, meaning, action, intentional, and dialectical. How each feature appeared in the reflections in this study is summarized below.

In both years, the reflections were anchored by an *experience* with a recent assignment (problems set, exam). In 2016, students reflected on an assignment that had been graded and returned;

so the experience was typically two weeks old. In 2017, students reflected on actions they had taken within the past week with regards to an assignment that was typically due the day of the reflection. Focusing on more recent experiences may have been easier for students to recall the details of their learning experiences.

In both 2016 and 2017, the *feature* of the experience the students were intended to focus on was the LOs with which they had the most difficulty. In 2016, the open-ended prompts did not always make students aware that focusing on their ability with individual LOs was desirable; they tended to focus on more general study skills (e.g. reading instructions carefully) [28]. In 2017, the embedding of self-ratings of each recently experienced LO in each reflection brought out this feature of the experience.

Students need a *lens* or knowledge to interpret or make meaning of the experience. In 2016, this may have been an issue. In end-of-semester surveys regarding students use of LO resources and feedback, students often wrote about not understanding what was meant by each LO [28]. In 2017, each LO was further detailed with a list of evidence of what proficiency with the LO means. As students had access to the LO lists during assignment and exam preparation, it is likely that students had greater knowledge of the goals for their intended learning than in 2016.

In 2016, due to the lack of a clear feature or lens, students had little structure for making *meaning* of their experience that would lead them to consider more effective ways of achieving LOs they found difficult. This may have led to them having a poor ability to identify an appropriate *action* or specific way to move forward [28]. It is unclear at this time if the prompt for action in 2017 elicited more LO-specific ideas. This will be explored in more detail using future reflections in which students can pick from a list of potential actions they could take. A list will be provided as there is some initial evidence that students are not aware of the many options available to them for improving their learning.

The few reflections conducted in 2016 did not convey a need for *intentional* (explicit rather than implicit) reflection. The weekly implementation of the 2017 reflections was a step in the direction of intentionality.

The *dialectical* feature of reflections was missing from the reflections implemented in this study. As stated in the methods section, these reflections were not intended to push students towards exploring their experience from different lenses or revising their lens. Rather, these reflections were viewed as an initiation into reflection that might prepare them to engage in less structured or deeper reflections in the future.

All told, the 2017 reflections were more theoretically sound than those used in 2016 and therefore more likely to have engaged students in improved self-regulated learning behaviors such as accessing rubrics to view their feedback. However, there was room for improvement in terms of the lens, action, and dialectical opportunity.

LIMITATIONS

This study has some limitations. First, only a few sections of a very large course were included. As such, results may have been influenced by the instructor or other section-specific factors. In this case, the instructor in the reflection section was deeply involved in the design of the SBG system. Expanding the study to include a greater number of reflection and non-reflection sections could account for section factors.

Second, the study was conducted in an FYE course with a low failure rate. What grade on an assignment is the threshold for inspiring a student to look at their feedback? Future work could investigate the feedback access behaviors of students in a course with a different final course grades distribution.

Third, this study does not account for the feedback access behaviors of students who might have had prior experience with a fully developed grading system with rich feedback, such as described in [39]. The prerequisite course for the FYE course in this study was just starting to transition to an SBG system in Fall 2017. Future work might look at students' access behaviors across multiple SBG courses in a plan of study.

Fourth, it is also unknown if the same results could have been obtained by reducing the number and structure of the 2017 reflections as the semester progressed. This would have placed more burden on the students to take up the mantel of intentional reflection, ideally what should be expected of self-regulated learners. Future studies might investigate how and when the structure of the reflections could be altered to transition self-reflection from the course to the students.

Fifth, the reflections used in this study did not expect students to explore their learning experiences and actions from multiple perspectives. Future studies might look at how and when greater dialectical opportunity should be introduced.

Finally, this study only examined students' access to feedback. No attempt was made to assess the time students spent reading their feedback, their interpretation of the feedback, or changes in their learning that might have resulted from reading feedback. Multiple opportunities exist for research in this space.

CONCLUSION

In this study, a large percentage of first-year engineering students did not review their standards-based feedback as delivered through the rubrics of a learning management system. Such a lack of access to feedback limits students' opportunities to learn, particularly when a system of grading is being employed that provides rich feedback, such as with SBG. The pairing of structured and frequent reflections with a well-developed SBG system was shown in this study to have potential to improve students' access to their feedback.

ACKNOWLEDGMENT

This work was made possible by a grant from the National Science Foundation (NSF DUE 1503794 and NSF IIS 1552288). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] Hattie, J., and H. Timperley, "The power of feedback," *Review of Educational Research*, vol. 77, no. 1, pp. 81–113, 2007. doi: 10.3102/003465430298487

[2] Bellon, J., E. Bellon, and M. A. Blank. *Teaching From a Research Knowledge Base: A Development and Renewal Process*. New York: Macmillan Publishing Company, 1992.

[3] Yorke, M, "Academic failure: a retrospective view from non-completing students," in *Failing Students in Higher Education*, M. Peelo and T. Wareham, Eds. Maindenhead, UK: SRHE/Open University Press, 2002.

[4] Shute, V. J., "Focus on formative feedback," *Review of Educational Research*, vol. 78, no. 1, pp. 153–189, 2008. doi: 10.3102/0034654307313795

[5] Black, P., and D. Wiliam, "Assessment and classroom learning," *Assessment in Education*, vol. 5, no. 1, pp. 7–74, 1998. doi: 10.1080/0969595980050102

[6] Alverno College Faculty, *Feedback is Teaching*. Milwaukee, WI: Alverno College Institute, 2015.

[7] Price, M., M. B. O'Donovan, C. Rust, and J. Carroll, "Assessment standards: a manifesto for change", *Brookes eJournal of Learning and Teaching*, vol. 2, no. 3, pp. 1-2, 2008.

[8] Reeves, D. B., *Making Standards Work: How to Implement Standards-Based Assessments in the Classroom, School, and District*, 3rd ed. Englewood, CO: Advanced Learning Press, 2003.

[9] Marzano, R. J., *Formative Assessment and Standards-Based Grading*. Bloomington, IN: Marzano Research Laboratory, 2010.

[10] Gooblar, D., "Pedagogy Unbound: Getting them to read our comments," ChronicleVitae, https://chroniclevitae.com/news/1129-getting-them-to-read-our-comments, Sept. 23, 2015.

[11] Diefes-Dux, H. A., "Student self-reported use of standards-based grading feedback and resources," in Proc. Research in Engineering Education Symposium, Bogota, Colombia, 2017.

[12] Zimmerman, B.J. "Attainment of self-regulation: A social cognitive perspective," in *Handbook of Self-Regulation*, M. Boekaerts, P.R. Pintrich, and M. Zeidner, Eds. San Diego, CA: Academic Press, 2000, pp. 13-39.

[13] Zimmerman, B. J., "Becoming a self-regulated learner: an overview," *Theory Into Practice*, vol. 41, no. 2, pp. 64-70, 2002. doi: 10.1207/s15430421tip4102_2

[14] Moon J. *Reflection in Learning and Professional Development*. London, UK: Kogan Page, 1999.

[15] Turns, J. A., B. Sattler, K. Yasuhara, J. L. Borgford-Parnell, and C. J. Atman, "Integrating reflection into engineering education," in Proc. American Society for Engineering Education Conference & Exposition, Indianapolis, IN, 2014.

[16] Turns, J. A, K. E. Shroyer, T. L. Lovins, and C. J. Atman, "Understanding reflection activities broadly," in Proc. American Society for Engineering Education Conference & Exposition, Columbus, OH, 2017.

[17] Dewey, J. *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*. Boston, MA: D.C. Heath & Co., 1933.

[18] Kolb, D. A. *Experiential learning: Experience as the source of learning and development*, 2nd ed. Upper Saddle River, NJ: Pearson Education, Inc, 2015.

[19] Schön, D. A. *The Reflective Practitioner: How Professionals Think in Action*. New York, NY: Basic Books, 1983.

[20] Sepp, L., M. Orand, J. Turns, L. Thomas, B. Sattler, B., and C. J. Atman, "On an upward trend: Reflection in engineering education," in Proc. American Society for Engineering Education Conference & Exposition, Seattle, WA, June 2015.

[21] Kisselburgh, L. G., J. L. Hess, C. B. Zoltowski, J. Beever, and A. O. Brightman, "Assessing a scaffolded, interactive, and reflective analysis framework for developing ethical reasoning in engineering students," in Proc. American Society for Engineering Education Conference & Exposition, New Orleans, LA, 2016.

[22] Hendricks, D. G., K. Yasuhara., and A. C. Taylor, "Work in progress: enhancing student leadership competencies through reflection," in Proc. American Society for Engineering Education Conference & Exposition, New Orleans, LA, 2016.

[23] Chenette, H. C. S. and T. Ribera, "Prediction and reflection activities in a chemical engineering course: fundamentals of heat and mass transfer," in Proc. American Society for Engineering Education Conference & Exposition, New Orleans, LA, 2016.

[24] Moon, J. "Getting the measure of reflection: considering matters of definition and depth," *Journal of Radiotherapy and Practice*, vol. 6, no. 4, pp. 191-200, 2007.

[25] Wilson, R. C., "Improving faculty teaching: effective use of student evaluations and consultants," *Journal of Higher Education*, vol. 57, no. 2., pp. 196-211, 1986.

[26] Angelo, T. A. and K. P. Cross, *Classroom Assessment Techniques: A Handbook for College Teachers*. San Francisco, CA: Jossey-Bass, 1993.

[27] Felder, R. M. and Brent, R. *Teaching and Learning STEM: A Practical Guide*. San Francisco, CA: Jossey-Bass, 2016.

[28] Diefes-Dux, H. A., "Student reflections on standards-based graded," in Proc. 46th ASEE/IEEE Frontiers in Education Conference, Erie, PA, 2016.

[29] Hatton, N. and Smith, D., "Reflection in teacher education: towards definition and implementation," *Teacher & Teacher Education*, vol. 11, no. 1, pp. 33-49, 1995.

[30] Ohland, M. W., M. L. Loughry, D. J. Woehr, L. G. Bullard, R. M. Felder, C. J. Finelli, R. A. Layton, H. R. Pomeranz, and D. G. Schmucker, "The comprehensive assessment of team member effectiveness: development of a behaviorally anchored rating scale for self- and peer evaluation." *Academy of Management Learning & Education*, vol. 11, no. 4, pp. 609-630, 2012. doi: 10.5465/amle.2010.0177.

[31] Dawson, P. "Assessment rubrics: towards clearer and more replicable design, research and practice," *Assessment & Evaluation in Higher Education*, vol. 42, no. 3, pp. 347-360, 2017. doi: 10.1080/02602938.2015.1111294.

[32] R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[33] Gardner, W., E. P. Mulvey, and E. C. Shaw, "Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models," *Psychological Bulletin*, vol. 118, no. 3, pp. 392-404, 1995.

[34] White, G. C., and R. R. Bennetts, "Analysis of frequency count data using the negative binomial distribution," *Ecology*, vol. 77, no. 8, pp. 2549-2557, 1996. doi: 10.2307/2265753.

[35] Bissell, A. F., "A negative binomial model with varying element sizes." *Biometrike*, vol. 52, no. 2, pp. 435-441, 1972.

[36] Cook, R. D., and Weisberg, S., *Residuals and Influence in Regression.* New York, NY: Chapman and Hall, 1982.

[37] Williams, D. A. "Generalized linear model diagnostics using the deviance and single case deletions," *Applied Statistics*, vol. 36, no. 2, pp. 181-191, 1987.

[38] Hicks, N. M. and H. A. Diefes-Dux, "Grader consistency in using standards-based rubrics," in Proc. American Society for Engineering Education Conference & Exposition, Columbus, OH, 2017.

[39] Diefes-Dux, H. A. and H. Ebrahiminejad, "Standards-Based Grading derived data to monitor grading and student learning," in Proc. American Society for Engineering Education Conference & Exposition, Salt Lake City, UT, 2018.