University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Biological Systems Engineering: Papers and Publications | Biological Systems Engineering

4-2022

Reflection Types and Students' Viewing of Feedback in a First-Year Engineering Course Using Standards-Based Grading

Heidi A. Diefes-Dux

Laura M. Cruz Castro

Follow this and additional works at: https://digitalcommons.unl.edu/biosysengfacpub

Part of the Bioresource and Agricultural Engineering Commons, Educational Methods Commons, Environmental Engineering Commons, Higher Education and Teaching Commons, Other Civil and Environmental Engineering Commons, Scholarship of Teaching and Learning Commons, and the Science and Mathematics Education Commons

This Article is brought to you for free and open access by the Biological Systems Engineering at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Biological Systems Engineering: Papers and Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Published in *Journal of Engineering Education* 111:2 (April 2022), pp. 283–307; doi: 10.1002/jee.20452
Copyright © 2022 American Society for Engineering Education. Published by Wiley. Used by permission.
Submitted August 7, 2020; revised September 12, 2021; accepted November 9, 2021.

Reflection Types and Students' Viewing of Feedback in a First-Year Engineering Course Using Standards-Based Grading

Heidi A. Diefes-Dux[1] and Laura M. Cruz Castro[2]

1. Department of Biological Systems Engineering, University of Nebraska–Lincoln, Lincoln, Nebraska, USA
2. School of Engineering Education, Purdue University, West Lafayette, Indiana, USA

Corresponding author – Heidi A. Diefes-Dux, Department of Biological Systems Engineering, University of Nebraska–Lincoln, East Campus, 210 L. W. Chase Hall, Lincoln, NE 68583-0726, USA, email heidi.diefes-dux@unl.edu

Abstract

Background: Feedback is one of the most powerful and essential tools for learning and assessment, particularly when it provides the information necessary to close an existing gap between actual and reference levels of performance. The literature on feedback has primarily focused on addressing strategies for providing effective feedback rather than aspects of students' readiness to engage with feedback. *Purpose/Hypothesis:* This study investigated whether reflection, as a routine pedagogical intervention grounded in self-regulated learning theory, promotes the frequency with which students view feedback. *Design/Method:* A quasi-experimental design was employed to examine the relationship between the use of four different reflection types, as well as no reflection, and students' feedback viewing behaviors in a first-year engineering course that used standards-based grading. Clickstream data were gathered through the learning management system to count the number of times students viewed feedback. The number of feedback views was compared by reflection type using descriptive statistics and a generalized linear model; weekly feedback viewing patterns were examined using time-descriptive graphs and the time-series cluster analysis. *Results:* Findings suggest reflection has the potential to increase the frequency of feedback views. Reflection not only had a positive and significant effect on the number of times students viewed feedback but also resulted in less variability between course sections and instructors when structured reflections made explicit references to feedback. *Conclusions:* Students need feedback to learn effectively, but many do not view feedback without formal prompting. The authors recommend instructors consistently administer reflections that include explicit pointers to feedback throughout the semester.

Keywords: assessment, feedback, first-year engineering, quantitative, reflection, self-regulated learning

1 Introduction

Formative feedback is defined as "information communicated to the learner that is intended to modify his or her thinking or behavior for the purpose of improving learning" (Shute, 2008, p. 154). Substantial evidence indicates that assessments that provide such feedback improve learning (Black & Wiliam, 1998), and feedback is "arguably the most important part of the assessment process" (Price et al., 2010, p. 277). What makes formative feedback critical to the learning process is its many roles, including "correction, reinforcement, forensic diagnosis, benchmarking, and longitudinal development (feed-forward)" (Price et al., 2010, p. 278). Ultimately, the power of formative feedback lies in a recognition of its ability to not only relay information about the quality of student work but also to prompt students' actions to close the gap between actual and reference levels of performance (Sadler, 1989).

Much effort has gone into identifying effective feedback strategies (e.g., Bangert-Drowns et al., 1991; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Shute, 2008). Good feedback practices include giving prompt feedback (Chickering & Gamson, 1987) and providing feedback that is criterion-referenced, tangible and transparent, actionable, user-friendly, ongoing, and consistent (Wiggins, 2012).

Establishing a grading system designed to employ these effective feedback strategies requires time and effort, particularly in the case of criterion-referenced grading strategies, such as standards-based grading (SBG). Criterion-referenced grading strategies, in general, aim to align expectations for student learning and assessment of student work (Sadler, 2005). SBG achieves this alignment through clearly articulated and shared learning objectives which are employed explicitly in the assessment of student work (Marzano, 2010). To develop and implement an SBG system, considerable instructor concentration and coordination is required (Lee et al., 2018).

Time and effort to create effective grading strategies, like SBG, would feel more justified were it not for the persistent problem of students' lack of engagement with feedback (Price et al., 2010). Instructors question students' attention to feedback (e.g., Badir & O'Neill, 2017) or believe students do not read the feedback (Mulliner & Tucker, 2017), particularly when they see students not retrieving their graded work (Nicol, 2010). That inattention leaves instructors feeling both concerned and irritated (Winter & Dye, 2004), and evidence suggests that instructors' laments may be warranted (e.g., Mensink & King, 2020). Students who do not view feedback are missing significant fodder for forward feeding reflection on their learning that could lead to personal development (Quinton & Smallbone, 2010). Ultimately, feedback "can have no impact unless students engage with it" (Handley et al., 2011, p. 553), and that engagement starts with students viewing feedback.

More than two decades ago, Chickering and Ehrmann (1996) described the uses of technology for giving feedback and foreshadowed the potential for technology to facilitate assessment systems. Today, learning management systems (LMS) allow instructors to provide feedback to students outside of class meetings. These systems also document feedback for performance tracking by both instructors and students. These technological capabilities should empower students to view feedback in a timely manner for the purpose of reflection and action, and yet, there are myriad factors related to why students do not view feedback (Mensink & King, 2020). Prior work has shown that first-year engineering students in a course using SBG, left to their own ends, did not view feedback in an LMS with a frequency commensurate with the number of graded assignments (Diefes-Dux & Cruz Castro, 2018).

Given what is known about students' limited use of feedback, Handley et al. (2011) proposed researchers focus on students' engagement with feedback, including "descriptions of the phenomena of engagement in terms of the diversity of students' readiness-to-engage" (p. 553). In a similar vein, Carless et al. (2011) advocated for reconceptualizing feedback—moving away from one-way communication by instructors to students and toward an interactive dialogue wherein student self-regulation is at the core of the feedback process. Self-reflection, the act of comparing one's performance to a standard "and attributing causal significance to the results" (Zimmerman, 2002, p. 21), is a key activity within self-regulation and has the potential to prompt students to view feedback as a means of improving their learning. Prior work by Diefes-Dux and Cruz Castro (2018) showed indicators that structured reflection can help achieve this goal. Structured reflection refers to prompting that makes explicit the components of critical reflection, including discussion of actions, thoughts on experiences, and development of plans (Aronson, 2011).

This study was an investigation into whether reflection, as a routine pedagogical intervention implemented in a course, encourages students to view feedback more frequently as one indicator of readiness to engage with feedback. This work, focused on reflection and students' frequency of feedback views, was conducted in two offerings of a large required first-year engineering course that used SBG. Results of this study were intended to inform engineering instructors and researchers of effective approaches to increasing students' engagement with feedback through enhancing students' lifelong learning abilities to seek, interpret, plan, and act on feedback.

In relation to this study, the first author was an instructor for over 20 years in the first-year engineering course that was this study's setting and oftentimes played a coordinating and content development role. She designed and conducted research on the SBG system and associated grader training and was, therefore, wellacquainted with instructors' commitment when employing SBG and students' limited use of feedback. Therefore, the approach taken in this study was influenced by first-hand knowledge of the course and prior research results. The second author had a background in statistics and was a doctoral student in engineering education. Both authors are pragmatists when it comes to education research and stand in agreement with the importance of empirical evidence when recommending teaching practices to improve student learning and advance engineering education.

2. Literature Review

Three bodies of literature inform this work. First, the literature on self-regulated learning (SRL) provides the theoretical basis for the use, design, and timing of reflections to prompt feedback viewing. Second, the literature on SBG is used to describe the assessment system into which reflection was integrated. And finally, past studies of students' feedback viewing behaviors are used to clarify the problem being addressed in this study.

2.1. Self-regulated learning: Theoretical framework

This study is anchored in the theory of SRL, which entails cycling through three phases: forethought, performance, and self-reflection (Zimmerman, 2002). Forethought entails goal setting and planning. Performance includes being aware of one's learning processes (actions) at the moment. Self-reflection requires comparing one's self-observed performance to some standard and reacting to the learning experience. Engagement in these phases aligns with qualities of lifelong learning, wherein one enacts behaviors of a reflective practitioner (Schön, 1983) and has been recognized for its role in academic achievement (Zimmerman, 1990).

Reflection is concerned with self-examination and exploration of an issue, grounded in an experience, to yield meaning and potentially a change in perspective (Boyd & Fales, 1983). In other words, reflection is critically thinking about an experience for the purpose of moving one's learning forward (Quinton & Smallbone, 2010). Mutch (2003) contended that learning from feedback is effective only if classroom time is dedicated to reflection on feedback, and Laflen and Smith (2017) recommended the use of reflection to engage students with feedback.

In this study, the experience that grounded students' reflections on their learning was their work on course assignments, and engagement with feedback was seen as a means to make deeper meaning of the gap between one's perception of their performance and the standard one is trying to achieve. Handley et al. (2011) conceptualized students' readiness-to-engage with feedback as a boundary between the instructor and the students across which artifacts are exchanged and expectations for learning and engagement are established (Figure 1a). For a given assignment, a sequence of exchanges occurs that sets the stage for student engagement with feedback. First, the instructor sets expectations by posting an assignment (*A*). Students then generate expectations for feedback through their submitted work on the assignment (*S*), and the instructor responds to the submissions with feedback (*F*). Finally, and ideally, students view feedback and engage with it.

Students' engagement with feedback in Handley and colleagues' conceptualization of students' readiness-toengage with feedback (Handley et al., 2011) can be conceived of as one or more cycles of the phases of SRL as described by Zimmerman (2002) (Figure 1b). Engagement with feedback then takes the form of self-reflection (*R*), which spurs planning (*P*) for improved learning, which is followed by performance, or actions (AC), taken with greater awareness of expectations. In this study, when no reflections were specifically assigned to students upon their receipt of feedback, such cycles of the phases of SRL were seen as being informally engaged by the students. When reflections were assigned upon the submission of their work, students were set on a path to formally engage in the phases of SRL (Figure 1c). In these instances, students completed a reflection in class, as recommended by Mutch (2003), immediately after an assignment was due, allowing them to reflect on their performance. The timing of reflection was intended, in part, to make plans to attend to feedback habitual.

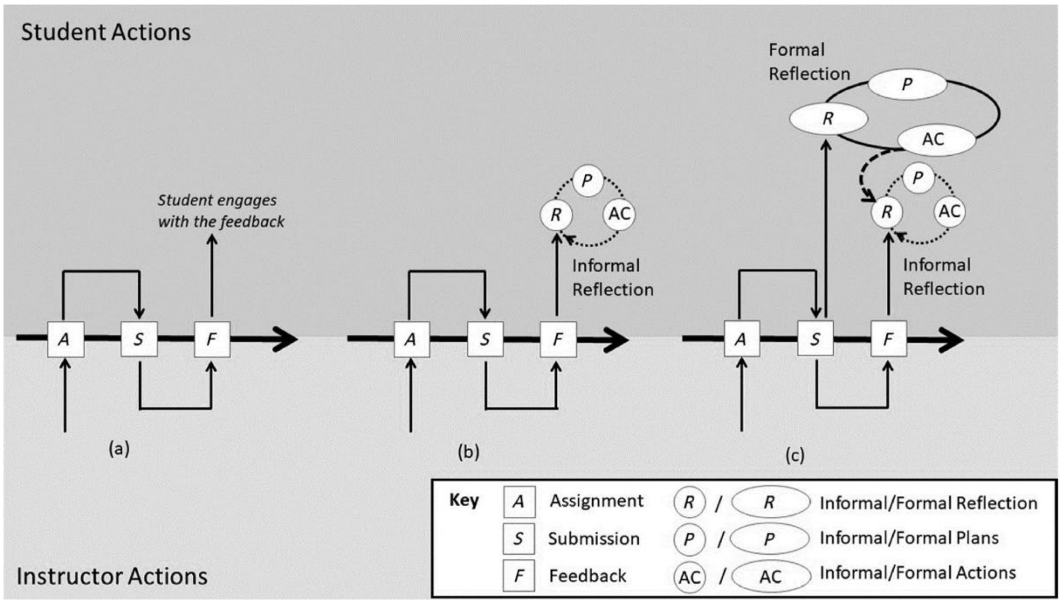

Figure 1. Students' readiness-to-engage with feedback: (a) as conceived by Handley et al. (2011), (b) with informal reflection as part of self-regulated learning (SRL) as conceived by Zimmerman (2002), and (c) with formal reflection driving informal reflection as used in this study.

2.2. *Standards-based grading*

SRL is predicated on the availability of clear standards (expectations for learning) to which one can refer when assessing one's learning. Criterion-referenced grading strategies provide transparency about the expectations for learning by using stated expectations to assess student performance. These expectations are communicated across the boundary between instructor and students through assignments and assessment (Figure 1).

SBG is a criterion-referenced form of assessment (Guskey, 2001; Heywood, 2014). The clear articulation of expectations in the form of learning objectives, the alignment of curriculum to those learning objectives, and the rich feedback based on those learning objectives that comprise an SBG system make this form of grading transparent and meaningful for students, instructors, and other stakeholders (e.g., Boud, 2017; Sadler, 2005; Scriffiny, 2008). In a discussion of the benefits and challenges of SBG in relation to Nilson's (2014) 15 criteria for a good grading system, Selbach-Allen et al. (2020) found SBG to rate favorably, with the rewards outweighing the challenges.

SBG is increasingly being used in engineering courses, from foundational courses, such as fluid mechanics (Post, 2014), thermodynamics (Mendez, 2018), and signals and systems

(Wierer, 2019) to engineering design (Atwood et al., 2014). Reasons to adopt SBG include the desire to focus students' attention on learning rather than grades (Mendez, 2018), improve the specificity of feedback to students (Wierer, 2019), and enable students to retest on particular learning objectives to show mastery (Post, 2014). Each of these reasons for SBG adoption requires that students actively view feedback and engage in self-assessment of their performance with regards to the learning objectives. Only then can students effectively plan and act to improve their learning. SBG was a critical component of this study as it provided the standards upon which students could reflect and make meaning of the gap between their learning and expectations for their learning.

2.3. Access to feedback

Gibbs and Simpson (2005) wrote, "It is not inevitable that students will read and pay attention to feedback even when that feedback is lovingly crafted and provided promptly" (p. 20). Empirical evidence to substantiate this notion is limited in general and particularly in engineering. The dearth of empirical evidence is likely because, prior to the data analytics capabilities of LMSs, studies were limited to investigations of faculty and student perceptions of students' handling of feedback. For example, Winter and Dye (2005) conducted a survey study of instructors in the schools of sport, performance arts and leisure, and education in the United Kingdom in which they asked instructors about uncollected student work. Instructors who responded to the survey estimated that more than 20% of graded work remained uncollected by students. Mulliner and Tucker (2017), in a survey study of staff and students in a school of the built environment in the United Kingdom, found that only 38% of instructors agreed that students always accessed their marked feedback, and only 35% agreed students always read their qualitative feedback, although 96% and 93% percent of the students agreed that they always viewed feedback or always read feedback, respectively. Diefes-Dux (2019), in a survey of end-of-semester first-year, engineering students, conducted with students who did minimal reflection, found that only 25% of students agreed that they reviewed their performance on weekly assignments soon after feedback was available early in the semester, although that percentage increased to 63% for feedback on later assignments.

LMSs, have made it possible to directly count the number of times students view feedback using clickstream data (i.e., data that track website links a user has selected). Few studies have used these data analytics to provide concrete evidence of infrequent feedback viewing. Mensink and King (2020) found that 38% of assessment files containing feedback generated in a series of biology modules were not accessed. Diefes-Dux and Cruz Castro (2018) found that, without reflection, 22 out of 115 students in a first-year engineering course never accessed their LMS-delivered feedback. In fact, over 75% of the students viewed feedback fewer than five times, though there were 20 assignments with feedback. The current paucity of empirical evidence that students do not frequently view feedback is one gap in the literature that this study fills.

Multiple factors have been explored as possible reasons for students' lack of viewing of feedback in higher education. One possible reason students do not view feedback is the inconsistent quality of the feedback they received in the past (Evans, 2013).

Lack of physical ease of accessibility to feedback has also been identified as a barrier to students viewing feedback. LMSs potentially offer better access to feedback in terms of timeliness and storage for on-demand access. However, evidence suggests that significant numbers of students do not view feedback even via an LMS (e.g., Mensink & King, 2020). Students' viewing of feedback via an LMS has also been found to be moderated by their concomitant access to the overall assignment grade. Laflen and Smith (2017) found that students were significantly less inclined to view feedback on written assignments when grades were accessible independently of the feedback. Similarly, Mensink and King (2020) saw 42% of assessment files go unviewed when grades were reported outside of feedback files.

Feedback viewing has also been found to be impacted by time in the semester. Students were found to be less inclined to view feedback on later assignments than early assignments in the semester, suggesting a perceived utility of the feedback within the course (Laflen & Smith, 2017; Mensink & King, 2020).

Whether feedback is provided physically on assignments or via an LMS, a number of strategies for encouraging feedback viewing have been articulated (e.g., Laflen & Smith, 2017; Winter & Dye, 2005), including early course emphasis on feedback, particularly on low-stakes assignments, explanations of the purpose and value of feedback, and engagement of students in the feedback process. In the context of this study, SBG was consistently employed on all assignments via an LMS, which afforded periodic classroom conversations about students' performance on the learning objectives, thus highlighting the value of feedback.

2.4. Prior work and significance of this study

In summary, feedback is an important element of education, and students' habitual viewing of feedback is necessary for them to engage fully in SRL. Students' lack of feedback views diminishes the advantages of using criterion-referenced grading strategies, like SBG, which ultimately leads to reduced learning. Within a large required first-year engineering course, the first author sought to integrate SBG with formal reflection as a means of consistently engaging students with feedback by developing self-regulatory habits. It was anticipated that students who engaged in structured reflection, meaning reflection that incorporated references to the three phases of SRL as described by Zimmerman (2002), would on average view feedback more often and earlier in the semester than those that used little or no formal reflection.

The authors have shown that students who completed weekly structured reflection self-reported more timely review of feedback than students who did not complete reflections (Diefes-Dux & Cruz Castro, 2019b). Recognizing that self-report of the frequency of feedback views can be an overestimate (Ada & Stansfield, 2017), two prior studies began to look at LMS generated clickstream data to assess students' actual frequency of feedback views. First, Diefes-Dux and Cruz Castro (2018) found that students who were not instructed to complete weekly reflections infrequently monitored their performance. Modest single-prompt reflections (i.e., minute-papers) showed some improvement in self-monitoring, while structured reflection showed significant improvement with the average number of feedback views being close to the number of assignments, though the standard deviation was large. Second, Diefes-Dux and Cruz Castro's (2019a) first attempt to study time-series patterns in the frequency of feedback views showed indicators that students prompted to reflect had different monthly viewing patterns, as compared to students who were not prompted to reflect, and their viewing of feedback appeared to start early in the semester. The focus of the study in this paper was a more nuanced look at students' feedback viewing behaviors when no reflection and different types of reflection were used in conjunction with SBG. Results were expected to garner greater insight into the potential for reflection to influence students' feedback viewing behaviors, a necessary first step toward improving students' readiness to engage with feedback.

3. Research Questions

This work sought to determine whether distinct patterns of feedback viewing emerge in relation to students' engagement in one of four reflection types, as well as no reflection. In particular, the research questions of concern were the following:

> RQ1: How do the total number of feedback views across the semester differ based on reflection type, including no reflection?
>
> RQ2: How do weekly feedback viewing patterns across the semester differ based on reflection type, including no reflection?

4. Methods

Reflections of various types were implemented in a first-year engineering course during two semesters. All reflections were classified as reflection-on-action, meaning thinking on a past experience, as opposed to reflection-in-action, which is thinking during an experience (Schön, 1983). To investigate the potential relationship between reflection type and students' viewing of feedback, clickstream data collected via the course LMS were used to indicate instances of students opening feedback. Descriptive and inferential statistics were used to pursue the research questions.

4.1. Setting and participants

This study was set in a required first-year engineering course offered in the Spring of 2017 and 2018 at a Midwestern United States university. The course enrollment was large ($N(2017) = 1600$, $N(2018) = 1521$) with approximately 25% female, 63% White, and 18% international students (Table 1).

Table 1. First-year engineering student participants' self-reported demographics by reflection type

	2017 (N = 1600)			2018 (N = 1521)		
Demographics	RA[a]	RB	NR	RC	RD	NR
Gender						
Female	25.0%	23.1%	23.9%	25.7%	17.1%	28.1%
Male	71.7%	76.5%	75.0%	73.1%	81.1%	70.6%
Other	0.3%	0.0%	0.4%	0.7%	1.8%	0.2%
No response	3.0%	0.4%	0.8%	0.4%	0.0%	1.1%
Race/ethnicity						
Asian	22.4%	18.8%	23.1%	18.7%	21.2%	16.7%
Black	1.6%	2.6%	2.1%	2.4%	5.0%	1.3%
Hispanic	6.6%	9.0%	9.2%	11.5%	10.4%	9.1%
Native American	0.0%	0.0%	0.1%	0.3%	0.0%	0.2%
White	60.5%	64.5%	59.4%	60.4%	56.3%	66.0%
Declined	2.6%	1.3%	2.1%	3.0%	2.3%	2.4%
Other	3.3%	3.4%	3.2%	3.3%	5.0%	3.3%
No response	3.0%	0.4%	0.8%	0.4%	0.0%	1.1%
International						
No	78.0%	83.8%	77.8%	82.4%	86.0%	86.6%
Yes	19.1%	15.8%	21.4%	17.2%	14.0%	12.2%
No response	3.0%	0.4%	0.8%	0.4%	0.0%	1.1%
Total	19.0%	14.6%	66.4%	44.0%	14.6%	41.4%

[a]Reflection types: RA (focused on feedback), RB (minute-papers), RC (focused on learning strategies), RD (personalized by instructor), NR (no reflection)

This large course was divided into sections, where a section is a division of the overall course enrollment that pursues the course learning outcomes with its own unique set of students and its own instructor (or instructional team) for the duration of the semester. In Spring 2017 and 2018, this first-year engineering course was divided into 14 course sections, each with at most 120 students, who met for 110 min twice each week. Each section was assigned an instructional team consisting of the lead instructor, a graduate teaching assistant (TA), and five undergraduate teaching assistants. Some lead instructors and graduate teaching assistants were assigned to multiple sections (Table 2).

Table 2. First-year engineering course section descriptions

2017				2018			
Self-regulated learning strategy	Section	Instructor code	TA code	Self-regulated learning strategy	Section	Instructor code	TA code
Structured reflection focused on feedback (RA)	RA1	C	a	Structured reflection focused on learning strategies (RC)	RC1	E	d
	RA2	A	b		RC2	E	d
	RA3	B	a		RC3	B	j
					RC4	M	k
					RC5	K	c
					RC6	K	c
Minute-papers (RB)	RB1	L	b	Structured reflection personalized by instructor (RD)	RD1	L	M
	RB2	F	f		RD2	L	M
No reflection (NR)	NR1	D	c	No reflection (NR)	NR11	M	k
	NR2	D	g		NR12	D	f
	NR3	E	d		NR13	D	n
	NR4	H	d		NR14	D	N
	NR5	D	e		NR15	N	p
	NR7	G	g		NR16	P	P
	NR8	H	h				
	NR9	H	h				
	NR10	K	c				

Abbreviation: TA, graduate teaching assistant

The goals for this 16-week course were centered on developing students' abilities to (1) apply basic (MATLAB) programming concepts to engineering problems, (2) represent and interpret data, (3) develop mathematical models to solve an engineering problem, (4) function effectively as a member of a team, and (5) demonstrate habits of a professional engineer. To this end, all sections followed the same curriculum with common lecture materials and resources, assignments, and exams. In 2017, 24 course activities were graded with learning objective-based rubrics, including five exams, five project milestones, 12 problem sets, and two teaming-related assignments. In 2018, there were only 10 problem sets, and thus, 22 graded course activities. The primary instructional difference between sections was the use of allotted extra credit that amounted to 2% of the total course grade. Some sections used these points to engage students in self-regulation strategies to improve learning (versions designated RA#–RD# in Table 2). The remaining sections did not use reflection exercises (designated NR# in Table 2).

A well-developed SBG system (Diefes-Dux & Ebrahiminejad, 2018) was used in this course, meaning, tight alignment existed between the assessment strategy and the stated course learning objectives (LO); high transparency of learning expectations was provided; graders received training; and data generated from the SBG system were used in the short term to guide classroom instruction and in the long term to improve the course. Each graded activity was assessed based on students' proficiency with a series of relevant LOs for which specific evidence of performance was sought. Grading was primarily done through LO-based rubrics in an LMS, specifically Blackboard Learn™. LO ratings were available to students through Blackboard rubrics and could be accessed via the Blackboard gradebook. For any LO attempted that received less than a proficient rating, written problem- and student-specific feedback was provided to indicate missing or incorrect evidence, occasionally with suggestions on how to improve. Written feedback ranged from short phrases to two sentences per LO. Grading was done primarily by the undergraduate teaching assistants and supervised by the TA. In 2017, online grader training was launched to improve the reliability of the grading (Hicks, 2020; Hicks & Douglas, 2018).

4.2. 2017 reflections

4.2.1. Structured reflections focused on feedback (RA)

For the purpose of this research, structured reflection refers to the use of specific prompts to draw students' attention to their performance on course learning objectives and help them consider actions and plans to improve their learning. In 2017, weekly structured reflections were completed in three sections of the course (RA1–RA3 in Table 2) via Blackboard 13 times across the semester. Reflections occurred typically at the start of the first class each week when students had 2 or 3 days to review feedback on a previously submitted problem set and had just submitted the next problem set prior to the start of class.

The structured reflections were designed to prompt students' plans and actions by aligning with Zimmerman's (2002) three SRL phases and making specific references to the course learning objectives. Typically, a reflection consisted of four recurring open-ended prompts (Table 3) plus a series of closed-ended prompts that asked students to rate their abilities with specific learning objectives that had recently been practiced (see Diefes-Dux and Carberry (2019a, 2019b) for an analysis of students' open-ended responses). 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 for the self-assessment prompts were the following: 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. As shown in Table 3, each RA reflection prompt aligned to one of Zimmerman's (2002) self-regulation phases, and all three phases were represented in the reflection.

Table 3. Structured reflection prompts for RA, RC, and RD

Question number	Reflection prompts	Response type	Alignment to Zimmerman (2002)
2017 structured reflection focused on feedback (RA)			
Q1	I have gone on Blackboard and reviewed my feedback on PS#a. [If no, skip to Q4]	Yes/No	Performance
Q2	Based on your feedback, what actions do you still plan to take to improve your abilities? Refer to specific learning objectives and be specific about your planned actions.	Open-ended	Forethought
Q3	Based on your feedback, what actions have you already taken to improve your abilities? Refer to specific learning objectives and be specific about your actions.		Performance
Q4–Q(X^b-1)	[Rate abilities with recent learning objectives]	5-point scale	Self-reflection
QX	For those learning objectives that you are not able to do on your own, what do you plan to do to improve your abilities over the next week? Refer to specific learning objectives (LOs) and be specific about your planned actions.	Open-ended	Forethought
2018 structured reflection focused on learning strategies (RC)			
Q1	While completing PS# [or Exam#], what actions did you take to help you attain proficiency with the LOs? Check all that apply.	Select all that apply	Performance
Q2–Q(Y^b-3)	[Rate abilities with recent learning objectives]	5-point scale	Self-reflection
Q(Y-2)	Considering the evidence of proficiency for the LOs, what is going particularly well for you? Be specific.	Open-ended	Self-reflection
Q(Y-1)	Considering the evidence of proficiency for the LOs, what is particularly difficult for you? Be specific.	Open-ended	Self-reflection
Q(Y)	What do you plan to do to improve your proficiency with the LOs over the next week? Check all that apply.	Select all that apply	Forethought
2018 structured reflection personalized by instructor (RD)			
Q1	After reviewing your feedback, how confident are you feeling about meeting these overarching learning objectives?	5-point scale	Self-reflection
Q2–Q(Z^b-7)	[Rate abilities with recent learning objectives at evidence level]	5-point scale	Self-reflection
Q(Z-6)	You just submitted PS#. How confident do you feel about your submission? That is, how well do you feel you have shown you have understood the LOs and can apply them to the problems we asked you?	5-point scale	Self-reflection
Q(Z-5)	What resources did you use to help you submit the problem set?	Select all that apply	Performance
Q(Z-4)	If you did the reflection last week, are these the things you committed to trying at the end of your last reflection?	Multiple choice	Forethought
Q(Z-3)	Do you have any comments you want to make on your submission that you made today?	Open-ended	Self-reflection

Table 3. Structured reflection prompts for RA, RC, and RD, *continued*

Question number	Reflection prompts	Response type	Alignment to Zimmerman (2002)
Q(Z-2)	You are about to start working on PS#. What resources are you going to make use of to help you submit this PS?	Select all that apply	Forethought
Q(Z-1)	Regarding the PS that you just got feedback on . . . How well do you feel the grade matches your expectations on how well you did?	Multiple choice	Self-reflection
Q(Z)	I would like Prof. L to invite me to a tutorial on some of this content.	Multiple choice	Forethought

[a]PS# represents problem set number.

[b]X, Y, and Z represent the total number of questions on any given reflection, which was impacted by the number of learning objectives rated in any given week.

4.2.2. Minute-paper reflections (RB)

The students in two 2017 sections (RB1 and RB2) completed online minute papers in Weeks 1 through 11 via Blackboard. A minute paper is a short in-class writing activity that is a means of seeking feedback from students on their learning (e.g., Chizmar & Ostrosky, 1998). RB1 completed 10 and RB2 completed 14 minute-papers. Prompts were typical of the form "what did you learn today about topic X?" In three instances, around the return of the first two problem sets and the first exam, students were asked more explicitly about the feedback they received on an assignment and whether the LO-based assessment reflected their learning.

4.3. 2018 reflections

4.3.1 Structured reflections focused on learning strategies (RC)

In 2018, weekly structured reflections were completed in six sections of the course (RC1–RC6 in Table 2) via Blackboard. While the implementation strategy was the same as in 2017, the free-response prompts changed (Table 3). Based on students' written reflections in the previous year, faculty were concerned that students were unaware of the wide variety of learning strategies available. Thus, students were prompted to select plans and actions from a 14-item list developed specifically to raise students' awareness of learning strategies, including referring to the posted LOs list detailing evidence of proficiency, watching and taking notes on the online modules, asking questions of classmates, and fixing problems on the previous problem set. The first author also desired that the students engage in a more thoughtful self-assessment of their performance with the course content (i.e., the LOs) than was seen in their reflections in 2017, as per Diefes-Dux (2016). Thus, the open-ended prompts asked students to consider the evidence of their proficiency with the LOs and to state what was going particularlywell and what was particularly difficult. The modifications altered the alignment to Zimmerman's (2002) self-regulation phases (Table 3) and resulted in the RCs (implemented in 2018) being less specifically about attending to feedback when developing plans and monitoring actions than the RAs and more about cognitively processing their achievement of the LOs. In total, students completed 10 RC reflections inWeeks 2–9, 11–13, and 15.

4.3.2 Structured reflections personalized by instructor (RD)

Sections RD1 and RD2 in 2018 completed reflections similar to RC (Table 3). However, instructor L personalized and managed the implementation of their reflections separately from the RCs. The number of RD reflective prompts was greater than that for the RAs and RCs, but, overall, the RDs maintained many of the RC elements and alignment to all three of Zimmerman's (2002) self-regulation phases. Differences can be seen in the way questions were posed, the level of learning objective detail on which students were asked to reflect, and the nature of the open-ended prompts. The RDs engaged students in self-reflection from a couple of additional angles, asking students about their confidence in meeting the learning objectives and considering how well the assessment of their work matched their expectations. In addition, the prompts that asked students to rate their abilities with the LOs associated with the most recently submitted problem set were not at the LO level but at the level of the 2–10 detailed statements concerning evidence of proficiency that described each LO. For example, with regards to Learning Objective 04.01 *Create a script that adheres to programming standards*, students were asked to rate their ability with the evidence items, which included *Assign computed values to variables* and *Comment variables with descriptions and units*. As in the RAs and RCs, the students were asked to select from a list of options concerning the resources they used to complete the problem sets and what resources they were committed to using for the next problem set. The options presented to students were extended beyond those in the RAs and RCs to include resources that were unique to the instructor's sections (e.g., "Review Prof. [L's] slides about common mistakes people have made on these LOs in the past" and "Posted questions on Piazza") and in some cases were more conversational in nature (e.g., "Team or study group before we start it to come up with a strategy on how to solve it (P.S. – this is ok to do not catching you out on anything here)"). Additional prompts asked students whether they had additional comments about their most recent problem set submission or needed additional help from the instructor. The personalization of this reflection type goes a step beyond prompting SRL to create opportunities for interactive dialogue between the instructor and their students, as per Carless et al. (2011).

Nine RDs were conducted in Weeks 2–9 and 11. These reflections were accessible a few days before and after the reflection was officially done in class, though most students completed them in class. RD reflections were completed via Qualtrics rather than the course LMS.

4.4 Student engagement and reflection in the course

To clarify the organization of course assignments and how reflection was theoretically envisioned to occur, it is prudent to again consider Handley et al. (2011) depiction of the aspects of student engagement with assignment artifacts and Zimmerman's self-regulation phases (2002) in the context of the design of the engineering course and reflections used in this study. Figure 2 illustrates the ongoing exchange of course artifacts between the instructional team and the students along a timeline and the interlocking of one assignment's (i.e., problem set) sequence of exchanges with the next assignment. Each assignment's sequence of exchanges began with the posting of an assignment (A_n) by the instructional team. While students were working on that assignment, the feedback on a previous (F_{n-1}) assignment was released. The students submitted their work (S_n) on the current assignment (A_n) prior to the posting of the next assignment (A_{n+1}). This process was repeated each week of the semesterwith slight variations around exams and the transition to project work mid-semester.

As described in the Background section, it was expected that students would enact informal reflection upon the release of feedback (F_{n-1}) and process that feedback by reflecting on the feedback (R), planning to address the feedback (P), and taking some actions (AC). The formal reflections were placed such that upon an assignment submission (S_n), students were formally prompted to reflect on their learning.

While students in RA, RC, and RD sections in both years were asked to rate their perceived ability with the most recently practiced learning objectives and consider plans in light of those perceptions going forward, the self-report prompts concerning actions taken in these years were different (Table 3). The space of planning and actions each reflection type was intended to prompt is shown by the boxes in Figure 2. In 2017, the RA reflection prompts Q1–Q3 (Table 3) explicitly drew students' attention to the feedback on the previous assignment (F_{n-1}) and their past actions and future plans concerning that feedback. In 2018, the RD reflections had similar feedback attention drawing prompts, but the RC reflections asked students to recall more generally actions they had taken during the completion of the current assignment ($A_n \rightarrow S_n$) to improve their learning. The only RC reference to feedback was in the list of past and planned actions from which students could select (i.e., "Review(ed) my performance on the LOs for PS [Problem Set] XX as soon as it was released"). So, the RA and RD reflections more explicitly asked students to look at and process feedback than the RC reflections, which placed reviewing feedback among a larger set of actions students might take to improve learning.

In contrast, the minimal RB minute-paper reflections typically asked about students' perceptions of their learning in the moment rather than looking backward or forward. The students in the RB and all nonreflection (NR) sections were, in theory, reliant on conducting their own informal reflection.

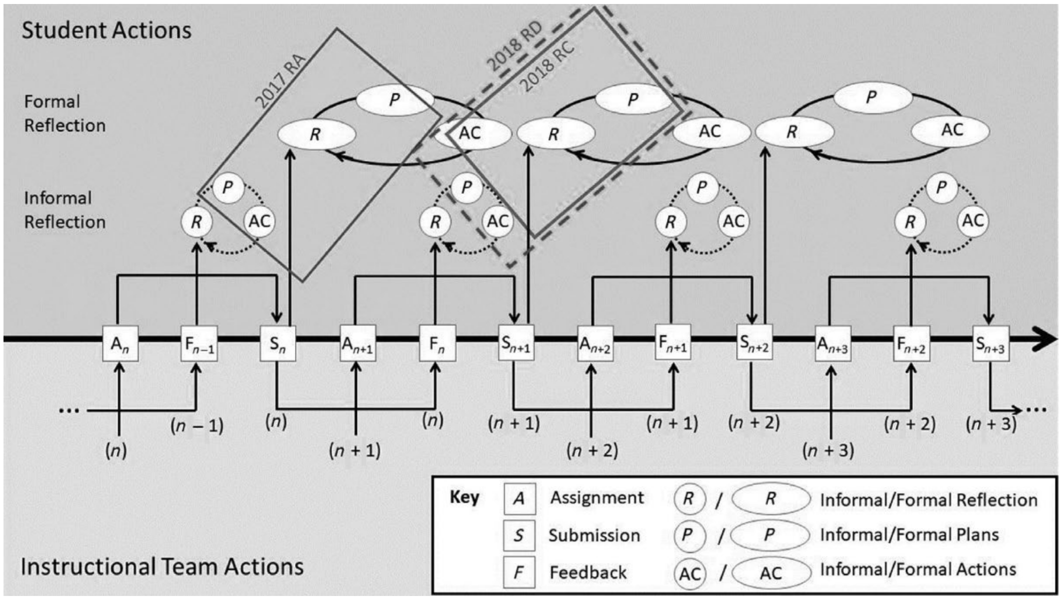

Figure 2. A contextualization of student engagement with assessment and feedback for this study based on Handley et al. (2011) and Zimmerman (2002).

4.5. Study design and data collection

A quasi-experimental, unbalanced study design was conducted, with two treatment (reflection) groups and one comparison (no reflection) group in each year. Each group included multiple but an unequal number of sections (Table 2).

The data for this study were students' clickstream data recorded as students navigated Blackboard to view feedback on the problem sets, exams, and project milestones. Clickstream data refer to the data collected by online platforms while users navigate the different features. These data usually include the object the user clicks, time of access, and session number. Rubrics showed the students how they performed on each LO assessed on an assignment and their written feedback on each LO. Data were recorded automatically by Blackboard each time a student accessed feedback by clicking View Rubric to see Rubric Details. A record of each view consisted of a student's ID, the timestamp of the access to the nearest minute, and a unique Blackboard login-session ID. It should be noted that each student's click to view feedback in a rubric is a call to the total Blackboard rubric database, not a specific rubric, as rubric data were not stored individually by rubric.

4.6. Data processing and analysis

The statistical software R (The R Foundation, n.d.) was used for all data processing and analysis. To start, observations that could create noise in the data were removed. The data of students who officially withdrew prior to the assignment of final grades were removed (20 students in 2017, 19 in 2018). In addition, any series of clicks by an individual student to View Rubric that occurred within 1 min were assumed to be a single view.

Multiple statistical approaches were used to analyze the data. For addressing the first research question, descriptive statistics and generalized linear models (GLMs) were utilized to gain insight into the overall number of feedback views by the group. For addressing the second research question, data visualizations were employed to reveal patterns in students' weekly feedback viewing behaviors by section. A time-series cluster analysis was then used to determine which sections displayed similar feedback viewing patterns.

4.6.1. Data analysis: Total feedback views (RQ1)

Descriptive statistics. Descriptive statistics concerning the total number of feedback views were conducted within the group in three ways: (1) including all students enrolled in the sections included in a group, (2) including only those students in a group who accessed feedback at least once, and (3) by sections included within a group.

Generalized linear models. A generalized linear model (GLM) and a generalized linear mixed model (GLMM) were used to model the effect of reflections on the total number of feedback views across the semester. Specifically, a zero-inflated negative binomial model (ZINB) and a posterior negative binomial mixed model (NBMM) were used, respectively. The R tool lme4 was used for the GLMs (Bates et al., 2015).

For both approaches, a chi-square test based on the log-likelihood, as well as the Akaike's information criteria (AIC) and the Bayesian information criteria (BIC) values, were used to select the best fit model possible with the available variables: type of reflection as the fixed effect and instructor and TAs as random effects.

The ZINB is a type of regression used when more zeros are present in the count data than can be predicted by standard negative binomial (NB) models (Faraway, 2016). This model can account for only fixed effects. It produces two sets of coefficients; the first set corresponds to the logistic model of the binary response of whether students viewed feedback at all. The second set corresponds to the truncated NB model, which accounts for the number of views for those who viewed feedback. The parameters for both the logistic and NB models are estimated with a quasi-maximum likelihood estimation (Lambert, 1992). The ZINB was chosen for three reasons: (1) the number of feedback views was a count, limiting the types of distributions that can be used to model feedback viewing behavior (Faraway, 2016), (2) the number of views was over-dispersed, meaning the variance was large compared to the mean (Lambert, 1992), and (3) the number of zeros that appeared in the data was large but important since they indicated whether students viewed feedback at all (Ridout et al., 2001). A Voung test was utilized to corroborate the importance of the zero-inflated part of the model (Vuong, 1989).

The variance between the two 2017 RB sections was unusually large (Table 4). The NBMM model allowed for the inclusion of random effects; specifically, instructors and TAs were treated as a hierarchical level to which the different types of reflection were applied. The inclusion of the random effects accounts for the variability that is encountered between sections with different instructors and TAs for the same type of reflection, making the estimation of the fixed effect (reflection type) more robust. As shown in Table 2, in 2018, the

RD group had only one instructor with two sections; therefore, the NBMM could not be used.

4.6.2. Data analysis: Weekly feedback views (RQ2)

Data visualization. Two different data visualizations were used to examine weekly feedback views. The first tracked each section's total average number of views per student immediately prior to the start of class each week. The second tracked the ratio of the total average number of views per student relative to the total number of assignments that had been returned to students with feedback prior to the start of class each week. The R tool ggplot2 was used for the data visualization (Wickham, 2016).

Time series cluster analysis. A time-series cluster analysis algorithm was used to determine whether the groups' feedback viewing patterns were distinct within each year. Cluster analysis is an unsupervised learning task that identifies homogeneous groups of observations. In this study, a cluster analysis was used to evaluate whether the variance between groups was large enough to be detected by a clustering algorithm. If the clustering algorithm places sections with the same type of reflection intervention in the same cluster, this means the patterns of feedback viewing for a given group were distinct. It should be noted that cluster analyses are descriptive techniques and therefore have no underlying assumptions. Like creating a bar chart of the data, it does not generate inferences, so the results are not generalizable to other populations.

In this study, a hierarchical clustering technique was used to identify the course sections included in each cluster. As opposed to other clustering techniques, the hierarchical technique provides a clear visualization of the clusters. In this clustering technique, objects (e.g., course sections) included in the clusters are added or removed repeatedly based on a calculated dissimilarity, which is measured as a distance between objects. Dissimilarity distance is a quantity that reflects the weakness of the relationship between objects in the clusters.

The Tsclust (Montero & Vilar, 2014) tool in R was used to compare clustering results obtained from four dissimilarity measures (i.e., dynamic time warping [DTW], dynamic time warping with L2 norm [Euclidean norm; DTW2], global alignment kernels [GAK], and shape-based distance [SBD]). The selection of the best dissimilarity measure and the number of clusters was based on the maximization of the silhouette index, which measures the relationship between cohesion and separation of the grouped observations (Rousseeuw, 1987). This step ensured that evidence of internal validity for the clustering pattern was found by the algorithm and that the number of clusters obtained represented the optimal separation between clusters and homogeneity (cohesion) within clusters.

By using the results from the optimization step, a time-series hierarchical cluster analysis was completed for each year. This analysis generated a dendrogram (i.e., a visualization of the clusters) for each year.

5. Results

In this section, the results of the quantitative analyses are presented as they pertain to each research question. The results of the analyses of students' total feedback views (RQ1) are followed by the analyses of their weekly patterns of feedback views across the semester (RQ2).

5.1. Total feedback views (RQ1)

5.1.1. Descriptive statistics

Table 4 shows the results of the descriptive statistical analysis of the total feedback views across the semester. These results are broken into three parts: those for all students enrolled in the course, those for students who viewed feedback one or more times, and those showing the variability of the sections within a group.

Table 4. Summary of students' feedback views

	2017				2018			
Reflection type	Ra[a]	RB	NR	Total	RC	RD	NR	Total
No. of sections	3	2	9	14	6	2	6	14
All students enrolled								
No. students	305	234	1061	1600	670	222	629	1521
Total views	6014	3663	10,238	19,915	9800	4035	8722	22,557
Average views (AV)	19.72	15.65	9.65	12.45	14.63	18.18	13.87	14.83
SD AV	13.14	11.88	9.88	11.62	11.10	12.40	10.51	11.14
SE AV	0.70	0.76	0.33	0.31	0.43	0.83	0.42	0.29
Ratio: Views to feedback[b]	0.82	0.65	0.40	0.52	0.67	0.83	0.63	0.67
% views in class	26%	51%	7%	20%	35%	40%	31%	34%
Students with one or more feedback views								
No. students	263	199	834	1296	613	206	577	1396
% students with no views	14%	15%	21%	19%	8.5%	7.2%	8.3%	8.2%
Average views	22.87	18.41	12.28	15.37	15.99	19.59	15.12	16.16
SD AV	11.35	10.76	9.59	11.04	10.62	11.75	10.07	10.67
SE AV	0.70	0.76	0.33	0.31	0.43	0.82	0.42	0.29
Ratio: Views to feedback[b]	0.95	0.76	0.51	0.64	0.72	0.89	0.69	0.74
Section variance								
SD AV	1.21	3.76	2.92	5.02	1.46	0.72	1.02	1.87
SE AV	0.70	2.66	0.97	1.34	0.59	0.51	0.42	0.50

[a]Reflection types: RA (focused on feedback), RB (minute-papers), RC (focused on learning strategies), RD (personalized by instructor), NR (no reflection)

[b]Average number of feedback views to the number of assignments with feedback.

In 2017, the average number of feedback views per student in the RA (focused on feedback) and RB (minutepapers) sections was considerably greater than that for students in the NR sections, with the RA section average being twice that of NR and the RB section being 50% greater. The ratio of the average number of views per student to the number of assessments with feedback the students could access was greatest for the RA section students (0.82), followed by the RB section students (0.65). This ratio for the NR sections (0.40) was comparatively low. Clicks to feedback occurred predominantly outside of class for the RA (75%) and NR sections (93%), while students in the RB sections accessed feedback equally in and out of class.

The variance in the number of views per student was lowest for the NR sections, which was likely due to the large number of NR students who never viewed feedback (21%) or rarely viewed feedback over the course of the semester. When the students who never viewed feedback were removed from the sample, the ratio of average views to available assignments with feedback increased for all groups, with RA section students nearly reaching parity (0.95). RB sections fared better than NR students at 0.76 and 0.51, respectively.

The variance among the two RB sections was quite high due to the total number of views for students in these sections being quite different (RB1 = 2158, RB2 = 1505) despite similar enrollments (118 and 116, respectively). The percentage of feedback views in class was also different for the RB sections (RB1 = 70%, RB2 = 41%). These were initial indicators of instructor variability that prompted the NBMM exploration of instructor and TA as random effects.

For 2018, the average number of feedback views per student in the RC (focused on learning strategies) and RD (personalized by instructor) sections was only slightly greater than that for students in the NR sections. The ratio of the average number of views to the number of available assignments with feedback was greatest for the RD students (0.83), followed by the RC students (0.67), which was similar to the NR section students (0.63). The percent of views that occurred in class was highest for the RD sections (40%).

Like 2017, when the number of students who never viewed feedback were removed from the 2018 data, the ratio of the average number of feedback views to the number of assignments with feedback available increased for all groups, with RD sections students reaching 0.89, and RC and NR students reaching 0.72 and 0.69, respectively. When comparing 2018 to 2017, there was a 44% increase in the average number of views per student in the NR sections. Regardless of section, in 2018, there were fewer students who did not view feedback (19% in 2017, 8.2% in 2018), and the difference between sections in this regard was small. Finally, from 2017 to 2018, there was a general increase in the number of views during class (20% in 2017, 34% in 2018).

5.1.2. Generalized linear models

The results for the ZINB for 2017 and 2018 are presented in Table 5, while the results for the NBMM for 2017 are presented in Table 6.

In 2017, the positive significant NB coefficient by reflection type (Table 5) indicates that there was a significantly greater number of feedback views when reflections were implemented than when they were not. The model predicts that there were, on average, 10.6 and 6.1 more feedback views per student for RA and RB students, respectively, than for NR students, which corresponds to the descriptive statistics results (Table 4). When modeling whether the students viewed feedback at all, both types of reflections were significant. The exp(logistic) values indicate that the probability of RA and RB reflection students not accessing the rubric system at all is 0.61 and 0.66, respectively, of the probability of NR section students (0.25). In other words, the model predicts that students in the reflection sections were less likely to not view feedback (15.2% of RA and 16.5% of RB students, values similar to those in Table 4).

Table 5. Summary of the zero-inflated negative binomial model (ZINB) for 2017 and 2018

ZINB 2017			ZINB 2018		
Variable	NB[a] (*SE*)	*exp*(logistic)	Variable	NB (*SE*)	*exp*(logistic)
Intercept	2.49 (0.025)***	0.25***	Intercept	2.70 (0.03)***	0.08
RA[b]	0.63 (0.05)***	0.61**	RC	0.06 (0.04)	1.05
RB	0.41 (0.05)***	0.66*	RD	0.26 (0.06)***	0.90
Theta[c]	2.33		Theta	2.26	
Log-likelihood[d]	–5405		Log-likelihood	–5547	

[a]NB refers to the estimated negative binomial model parameters.
[b]Reflection types: RA (focused on feedback), RB (minute-papers), RC (focused on learning strategies), RD (personalized by instructor).
[c]Theta is the estimated over-dispersion parameter used for testing the use of a Poisson linear model.
[d]Log-likelihood estimation serves as an indicator of the goodness of fit of the model when compared to other models.
*p < .01, **p < .001, ***p < .0001

Table 6. Summary of posterior negative binomial mixed model (NBMM) for 2017

Variable	NB[a] (SE)	Random effects variable	Variance (SD)
Intercept	2.52 (0.11)***	Instructor	0.041 (0.20)
Ra[b]	0.58 (0.19)***	TA	0.015 (0.12)
RB	0.36 (0.20)		
Theta[c]	2.53		
Log-likelihood[d]	–4612		

Abbreviation: TA, graduate teaching assistant.
[a]NB refers to the estimated negative binomial model parameters.
[b]Reflection types: RA (focused on feedback), RB (minute-papers).
[c]Theta is the estimated over-dispersion parameter used for testing the use of a Poisson linear model.
[d]Log-likelihood estimation serves as an indicator of the goodness of fit of the model when compared to other models.
****p* < .0001

For 2018, the positive NB coefficients for the RC and RD sections indicate that students viewed feedback more frequently when reflections were implemented than when they were not. The coefficient for the RD sections was significant and indicates a 4.4 increase in the number of feedback views over the NR section students. The coefficient for the RC sections, however, was not significant. When modeling the students who did not view feedback at all, both types of reflections were not significant, meaning that the 2018 reflections were not a good predictor of the probability of not viewing feedback at all.

When accounting for instructors and TAs, the positive NB coefficient for the RB reflection type (Table 6) indicates that there were five more feedback views per student for the RB than for the NR sections. However, the significance of the coefficient was reduced when the instructor and TAs were accounted for (comparing the RB NB values from Tables 5 and 6), implying that part of the effect of RB in 2017 corresponded to the instructor and TA variability. In contrast, the significance of the effect of the RA reflection type is not changed from the ZINB to the NBMM models, indicating that the instructor and TAs were not significant in predicting the number of feedback views. For 2017, the variance estimates for the random effects, as well as the change in the coefficient for RB, are an indicator of the importance of accounting for the variance induced by the instructional teams.

5.2. Feedback viewing patterns (RQ2)

5.2.1. Data visualization

The total average number of views per student were plotted each week of the semester for 2017 (Figure 3) and 2018 (Figure 4). The reference line, labeled REF, represents the total number of assessments with feedback that was available to the students.

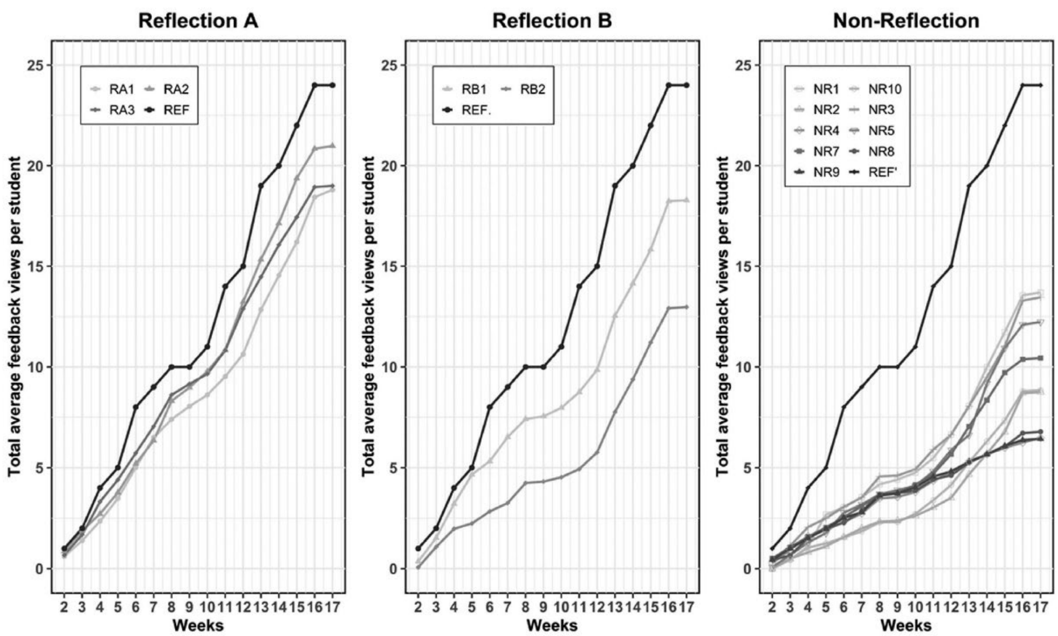

Figure 3. Total average number of feedback views per student on a weekly basis in 2017. REF is the total number of assignments with feedback available for access. Reflection types: RA (focused on feedback), RB (minute-papers), RC (focused on learning strategies), RD (personalized by instructor), NR (no reflection).

In both years, but particularly in 2018, when multiple feedback items were released within the same week (Weeks 6 in 2017; Week 8 in 2018), a gap appears between the sections' viewing behaviors and the reference line that may (as for the RA sections) or may not (as for all 2018 sections) be overcome in subsequent weeks (Figure 4). In 2018, this effect was seen again in Week 13, when multiple assignments were returned.

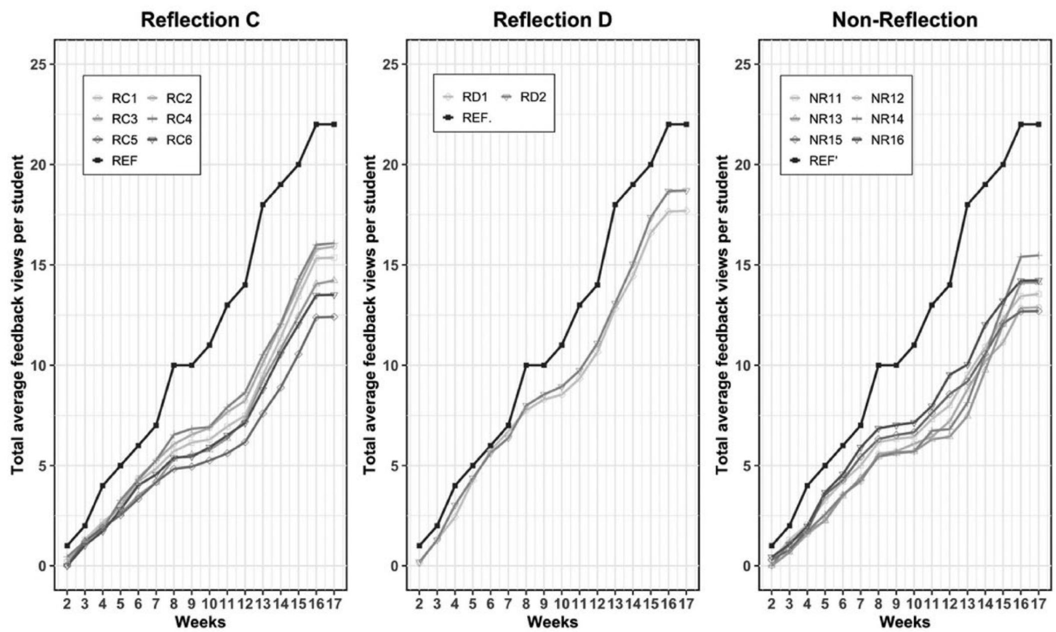

Figure 4. Total average number of feedback views per student on a weekly basis in 2018. REF is the total number of assignments with feedback available for access. Reflection types: RA (focused on feedback), RB (minute-papers), RC (focused on learning strategies), RD (personalized by instructor), NR (no reflection).

Figure 5 shows the ratio of the average number of feedback views to the number of available assignments with feedback each week. Sections using the same reflection type have similar line styles in these plots to draw attention to section patterns. If the students in a section were to, on average, view all feedback as it became available, the ratio would be consistently a value of 1 across the semester.

Figure 5 plots enable one to see whether sections in the same group track similarly to each other and differently from sections in other groups. In 2017, the feedback viewing behavior of RA sections was different than that of NR sections. The RA sections consistently viewed feedback at average rates hovering around 75% of what was available even from the start of the semester. The NR sections, however, had very low average rates of viewing in Week 2, rates in a band of 23%–45% up until Week 3, when some sections saw slightly increasing rates through to the end of the semester. The two RB sections were split in their behavior; RB1 tracked with the RA sections, while RB2 tracked with the NR sections.

In 2018, feedback viewing by students in the RD sections was consistently greater than the RC and NR sections by Week 4. In general, the viewing behaviors of the students in the RC and NR sections appeared indistinguishable. One particular contrast between sections is noteworthy; in Week 4, there was a noticeable drop in the students' viewing ratio for many RC and NR sections and one RD section.

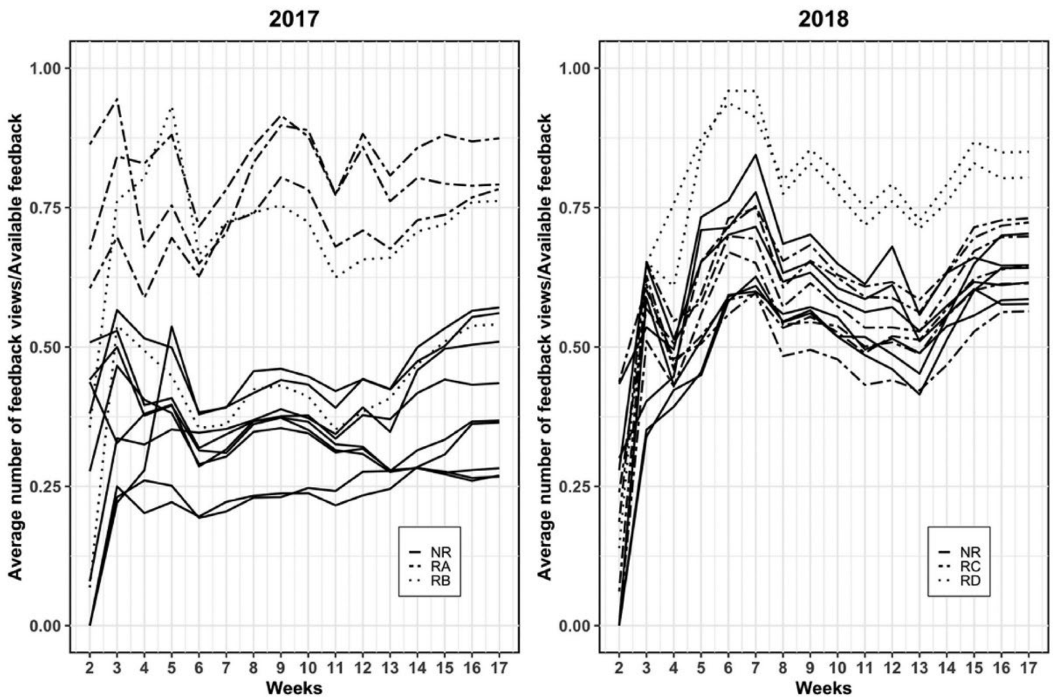

Figure 5. Ratio of average feedback views to the number of available assignments with feedback. Reflection types: RA (focused on feedback), RB (minute-papers), RC (focused on learning strategies), RD (personalized by instructor), NR (no reflection).

5.2.2. Time series cluster analysis

The GAK was found to optimize the silhouette index for both 2017 and 2018. However, a maximum silhouette index was obtained when the data were divided into three clusters in 2017 and two clusters in 2018 (Figure 6). Figure 7 shows how the sections clustered based on students' clickstream feedback viewing patterns feedback across the semester.

In 2017, the clustering algorithm detected three clusters: (1) all of the RA sections and RB1, (2) five NR sections (NR2, NR4, NR8, NR9, and NR10), and (3) the five remaining NR sections and RB2 (Figure 7). The RA sections were likely clustered due to their relatively larger average number of feedback views (Table 4), as well as viewing behavior patterns that more closely tracked with the availability of feedback than other sections (Figure 5). The sections in the second and third clusters were likely separated due to differences in the number of feedback views between Weeks 8 and 12, wherein cluster two sections, particularly NR4, NR8, and NR9, were least likely to view feedback in the last weeks of the semester (Figure 3).

In 2018, the feedback viewing behavior of the RD sections clustered separately from the RC and NR sections (Figure 5). The RD sections likely clustered separately due to the larger average number of feedback views (Table 4), as well as viewing behaviors that tracked more closely to the release of feedback (Figure 5).

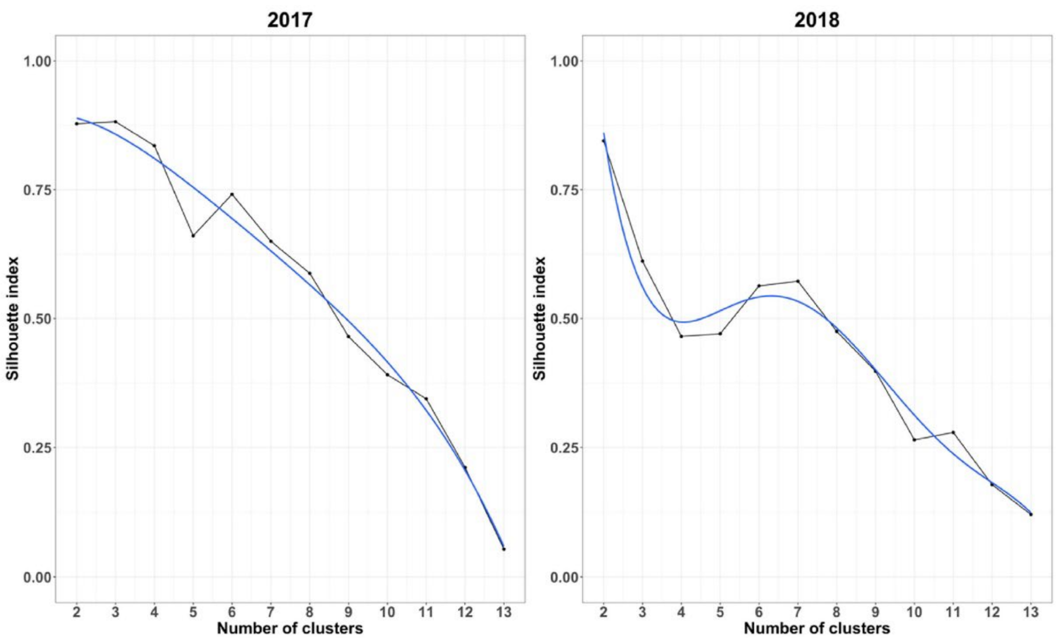

Figure 6. Optimization of number of clusters using global alignment kernels (GAK).

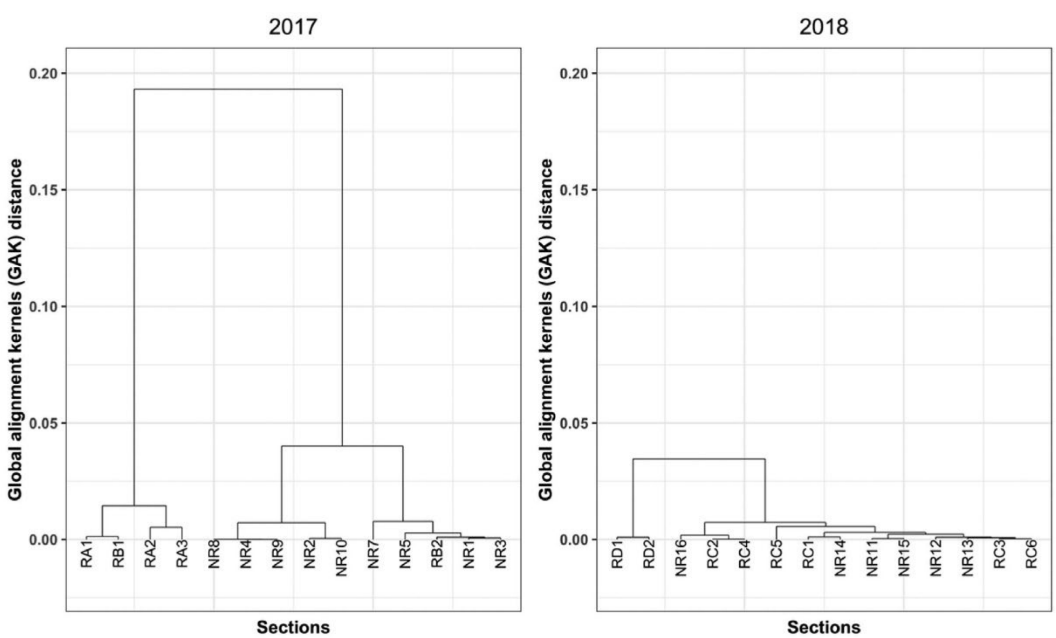

Figure 7. Dendrograms from the hierarchical time-series cluster analysis. Reflection types: RA (focused on feedback), RB (minutepapers), RC (focused on learning strategies), RD (personalized by instructor), NR (no reflection).

6. Discussion

Each of the two research questions is addressed below. For RQ1, which was concerned with differences in the total number of feedback views across the semester based on reflection type, the discussion elaborates on the results of the summary statistics, linear model, and participation in reflection, leading to the discovery of the confounding effect between reflection, instructor, and operational changes. For RQ2, which was concerned with differences in weekly feedback viewing patterns across the semester, the discussion focuses on the results obtained by time-descriptive graphs and the time-series cluster analysis. This discussion centers on specific points in time and trajectory patterns, offering additional information to explain the dips or increases in students' feedback views in specific time frames. For this purpose, additional information from the reflections, instruction, and course policy are discussed. Finally, the potential of the four types of reflection (RA, RB, RC, and RD) to improve students' feedback viewing behavior is presented.

6.1. Reflection and total feedback views (RQ1)

From the descriptive statistics, it was evident that students in all reflection sections, on average, viewed feedback more frequently than students in NR sections. Further, from the GLMs, it was evident that the RA (structured reflection focused on feedback) and RD (instructor personalized reflection on learning strategies) sections had a statistically significant effect on the total number of feedback views per student. The RA SRL (Zimmerman, 2002) aligned performance Q1 prompt, and the RD SRL aligned self-reflection Q1, and Q(Z-1) prompts that focused students' attention explicitly on feedback (Table 3) seem to be a likely explanation for why RA and RD students viewed feedback more often than NR students. However, it is interesting that many of the RA section feedback views occurred outside of class, perhaps suggesting that those students who were viewing feedback were planning ahead for the reflections each week and were therefore developing a habit of viewing feedback and engaging in some level of informal reflection. This response was envisioned for RA with the strategic placement of the reflection in the instructor-student exchange, placing it at the time an assignment was submitted (Figure 1c) when the students' experience with prior feedback and the most recent learning objectives were fresh and triggered for planning that could include viewing feedback. Lower RD reflection completion rates (~60% RD vs. ~90% RA) may not have inspired similar habits.

The feedback viewing behaviors of the other reflection sections and the change in the total number of feedback views of the NR (no reflection) sections from 2017 to 2018 require other considerations. To start, the total number of feedback views for RB1 is surprising as the reflections were unstructured and therefore did not explicitly guide students to critically reflect (Aronson, 2011) on each of the SRL phases of forethought, performance, and self-reflection (Zimmerman, 2002). Further, the two RB (minute-papers) sections had very different total feedback views. While students in both sections completed similar minute papers, the NBMM results indicated that the instructional team mattered. The high total number of views for RB1 students indicates some motivator was not present in RB2. Yet this motivator did not appear to differentially impact students' participation in reflection (~95% for both sections through Week 6) or students' performance on assignments. Finding that the instructor was a factor parallels that of Klobas and McGill (2010), who found that instructor involvement increased students' use of an LMS.

The instructor factor likely continued to have a very strong impact on the RD sections as the instructor of the RB1 and RD sections was the same. It might seem obvious to point to the more detailed and personal prompts in the RD reflection as the motivation for the greater number of RD students' feedback views. However, the completion rates of the RD reflections (~60%) were far lower than for the RC (structured reflection focused on learning strategies) reflections (~90%). The difference in these completion rates again points to something occurring in the RD sections, other than the reflections, that motivated students to view their feedback, perhaps some element of instructor rapport as may be seen in the last RD prompt that focused on personal communication with the instructor or the conversational tone of the reflection prompts (Table 3) and options that may invite greater dialogue around feedback (Carless et al., 2011) and learning strategies more generally.

The number of students with zero views and the low total number of feedback views for the NR sections in 2017 was consistent with the findings of other researchers (Laflen & Smith, 2017; Mensink & King, 2020; Winter & Dye, 2005). Yet, NR students viewed feedback more frequently in 2018. This increase may be explained by two 2018 course operational changes. First, students' learning objectives performance on assignments was processed in a weekly fashion in a way that instructors could share with their students (Diefes-Dux & Ebrahiminejad, 2018); course-wide, not section-level, SBG data were available in 2017. Instructor sharing of these SBG data with their students might explain the greater percentage of in-class feedback views across all 2018 sections as compared to 2017. Seeing their section's performance on the learning objectives may have motivated students to compare their own performance to that of the class. Even if section performance results were not explicitly shown in class, instructors' heightened awareness of student performance as presented in weekly instructional team meetings may have had an impact on the frequency with which instructors referred to the availability of feedback on assignments, thus prompting students to attend to feedback.

Second, although it has been noted by personal experiences in different settings that SBG is likely a more consistent grading system (Selbach-Allen et al., 2020), more rigorous grader training with the SBG system was instituted to improve the reliability and quality of feedback on assignments (Hicks, 2020; Hicks & Douglas, 2018). One indicator of improvement is the average number of characters in the graders' written feedback. The number of characters comprising the written feedback on learning objectives on which the instructional team was required to give feedback (due to less than proficient student performance) increased by about 25% from 2017 to 2018 (for the RC and NR sections, though interestingly not for the RD sections). Students' early interactions with feedback that set clear expectations and that they found valuable may have encouraged ongoing feedback viewing for the duration of the semester. Students may also have perceived that they benefited from ongoing feedback viewing in this course, as computer programming skills built on each other and attention to feedback across project milestones ensured success on the final project submission. So, higher quality feedback and students' perceived value and utility of the feedback to their course success (Lizzio & Wilson, 2008; Selbach-Allen et al., 2020) may have driven their feedback viewing.

6.2. Reflection and feedback viewing patterns (RQ2)

The visualizations and the cluster analysis provide evidence that the students' feedback viewing in the RA sections in 2017 and RD sections in 2018 were distinct from the other sections in their respective years. Students in these sections viewed feedback at a rate more comparable to when feedback was available than those in other sections. Further, the pace of RA student feedback viewing suggests a relatively timely engagement with feedback, as conceptualized in Figure 2 and self-reported by these students (Diefes-Dux & Cruz Castro, 2019a), and could indicate antecedents of SRL habits.

The fact that the feedback viewing pattern for RB1 (taught by instructor L) was clustered with the RA sections and not with RB2 (taught by instructor F) was again an indicator that the instructor factor was more significant than the minute-paper reflection type. The instructor factor was also likely the primary contributing factor to the distinctive feedback viewing behaviors of RD section students. That the feedback viewing patterns for the RC (structured reflection focused on learning strategies) sections were not found to be distinct from some of the NR sections seems to indicate the reflections were not particularly influential when it came to readying students to engage with feedback. The lack of explicit references to feedback in the reflection prompts aligned to SRL theory (Zimmerman, 2002) seems the likely explanation.

What is perhaps most interesting is that in Figure 5, in both years, there are peaks, dips, and trajectories in the students' feedback viewing patterns that are common to all or groups of sections. These commonalities may more clearly elucidate the story of how reflection did and did not promote feedback viewing.

First, consider the average number of feedback views in Week 2. The students in the RA sections viewed feedback more than students in any group in either year. It seems that RA section students were more aware that there was feedback to view, which may have been the result of RA instructors foreshadowing the implementation of reflections and extra credit in Week 1. The Week 2 zero views of feedback in a few NR sections in 2017 was due to a delay in the return of feedback which may have contributed to students' poor feedback viewing behaviors for the duration of the semester. In 2018, grades were released on schedule across all sections; however, few students viewed feedback in Week 2 regardless of whether reflection was done or not. It is likely that students were not made aware of the timing of the release of feedback or were unfamiliar with how to access feedback. While other researchers have suspected the latter to be an issue (Laflen & Smith, 2017), lack of familiarity with how to access feedback seems unlikely here as these students accessed feedback in a similar fashion in the prerequisite course. The spike in the average number of feedback views in Week 3 in both years likely coincides with the return of the second problem set, which was more difficult than the first, and a raised awareness that feedback was available.

Unfortunately, in Week 4, there is a pronounced drop in the average number of feedback views, particularly in 2018. One possible explanation for this drop is the distraction of exams in other courses, including math and chemistry. While there was a recovery in Week 5, in both years, it was not maintained in 2017, perhaps due to the timing of this course's first exam and the return of multiple feedback items. In 2018, the recovery from

Week 4 exams appeared sustained, but a Week 8 dip was seen following that semester's first exam and the return of multiple feedback items.

Mid-semester, RA sections appeared to gain momentum, whereas the frequency of feedback views for all other sections appeared to level off (2017) or steadily decline (2018). One conjecture is that the RA reflection prompts focused on feedback helped sustain and even promote the habit of viewing feedback.

The feedback viewing pattern is interrupted in both semesters and for all sections when the courses transitioned from problem set assignments to project milestones around Weeks 12 and 13. Following this transition, students in most sections, on average, showed an increase in feedback views, but with some exceptions. In 2017, three sections (NR4, NR8, and NR9) were found to have provided all written feedback to students on the physical copy of students' project milestone work, while their LO assessments still appeared in the LMS. This grading behavior continued to some extent in some NR sections in 2018. The tendency to provide feedback both on students' physical work and through the LMS may have diverted students' attention from the feedback on the LMS.

This discussion of the course details in relation to students' feedback viewing patterns points to the potential power of the RA reflection type and its consistent deployment in helping students maintain their readiness to engage with feedback. RA section students continued to view feedback at higher rates than almost all other sections despite disruptions within and outside of the course. While a similar effect is seen in sections taught by instructor L (RB1, RD1, and RD2), the elements of the instructor's strategy leading to students' high number of feedback views in total and across the semester were not identifiable and may have been independent of the use of reflections, creating a difficulty when discussing the potential of the RB and RD reflection types.

While the RC reflection type did not garner a significantly greater average number of feedback views than no reflection in 2018, this reflection type may not be without merit. A very preliminary look at the nature of students' responses to the RC open-ended response items revealed greater metacognitive monitoring and transfer than was elicited by the RA open-ended prompts (Stratman & Diefes-Dux, 2021). It may be that features of both the RA and RC reflection types may make for a more optimal reflection package in which some prompts facilitate the viewing of feedback and others draw students into deeper thinking about their learning. Future work will look more critically at the differences in students' open-ended responses.

7. Implications for Practice and Research

Within the greater scheme of cycles of students embarking on an assignment, submitting an assignment, reviewing feedback, and feed-forwarding into the next assignment, Hounsell et al. (2008) identified 13 different trouble-spots to help instructors identify and remedy guidance and feedback strategies. What was missing from their list, and is quite apparent from this study, is a lack of student feedback views. In this regard, there are three takeaways from this study. First, this study provided evidence that, when left to their own devices, students, in this case, first-year engineering students, may infrequently view feedback. This is significant for instructors engaged in grading reform, where the provision of meaningful formative feedback is intended to increase students' learning. If students are not opening feedback, they cannot benefit from it. If this habit of not viewing (or seeking) feedback becomes engrained, there may be downstream consequences for students' academic and professional success.

Second, this study provided evidence that an effective intervention to encourage students to view feedback is a periodic reflection with prompts that purposely draw students' attention to the feedback available to them. Evidence also suggested that reflections that attempt to raise students' awareness of a myriad of learning strategies at the same time may not have as significant an effect on the number of times students view their feedback. Meaning, perhaps not unexpectedly, that reflections that are focused on a single or limited number of purposes or clear learning goals (Aronson, 2011) are more likely to have the desired effect.

Third, this study provided evidence that the relationship between instructors and their students, the grading system, reflection, and feedback viewing behaviors are complex. While reflection, optimally designed and timed for effect, can elicit greater student viewing of feedback, there were indicators that suggest that instructor related factors (e.g., rapport with students, communication of expectations, engagement with assessment and feedback) and overall grading system design factors (e.g., quality and timing of feedback) can alter the effect. These potential mediators of the impact of reflection, as well as the effect of reflection type when using different grading systems, present opportunities for further study.

8. Limitations

Several limitations associated with this study limit the generalizability of findings. First, it was unknown from the clickstream data which specific assignment feedback a given student was viewing. So, the average number of views only enabled the researchers to think about whether the students, on average, were viewing the feedback at a rate commensurate with availability. Second, the participants were first-year engineering students. More advanced students, with or without reflection, may demonstrate different feedback viewing behaviors. Third, more than 60% of the participants were White, and more than 70% were males, making the sample size of marginalized groups (e.g., students of color and other gender identities) underrepresented in the sample. Fourth, the quasi-experimental design and some curriculum changes limited the direct comparison between years.

In addition, the setting of this study was complex. The large course enrollment and the size of the instructional team caused variation in the implementation of the course from section to section. This variation was seen, for example, in one section deciding to deliver feedback outside of the LMS. However, the variability was not as great as may be expected as there was strong central administration of the content, instruction schedule, grader training, and grading schedule, as well as instructor and staff commitment to a common first-year experience. Still, instructors brought their unique perspectives to content delivery, mentoring of the teaching assistants supporting their sections, and communication of expectations around reflection and feedback to students.

Finally, this study paired reflection with an SBG system. Different results might be found when utilizing a different grading strategy. It is possible to hypothesize that the potential for reflection to increase the number of feedback view is not anchored to a specific grading system. Nevertheless, it seems reasonable to expect that alignment between the reflection design and the elements of the grading system employed in a course is essential if the goal of reflection is to improve students' engagement with the grading system, and ultimately feedback.

9. Conclusion

In this study, weekly reflection guided by SRL theory was seen to have the potential to improve students' frequency of feedback views, particularly when prompts asked students to think about their past actions and future plans while keeping their perceived ability with the learning objectives and expectations, as conveyed through an SBG system, in mind. Generally, it was found that weekly reflection did increase the number of times students viewed feedback. When students completed reflections with prompts explicitly referring to feedback provided by the instructional team (RA), they accessed feedback significantly more often than those completing no reflections (NR1). Students who completed reflections, both of a minute-paper type (RB) and a highly structured type without highly explicit references to assignment feedback (RD), also viewed feedback more frequently than those that completed no reflections (NR1 and NR2, respectively), though the instructor effect was strong in these instances. For ensuring students are a step closer to being ready to engage with feedback across the semester, it is recommended that reflections be paired with the grading system, include explicit pointers to feedback, and be consistently administered to students throughout the semester. This recommendation is particularly relevant when using criterion-referenced grading systems, like SBG, that require significant effort to establish but have the potential to provide valuable benefits to students' learning and success.

Acknowledgments – This work was made possible by a grant from the National Science Foundation, NSF DUE 1503794 and NSF IIS 1552288. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation. The authors wish to express their gratitude to the first-year engineering instructional support team and the instructional teams.

ORCID
Heidi A. Diefes-Dux https://orcid.org/0000-0003-3635-1825
Laura M. Cruz Castro https://orcid.org/0000-0002-9331-090X

References

Ada, M. B., & Stansfield, M. (2017). *The potential of learning analytics in understanding students' engagement with their assessment feedback.* Paper presented at the IEEE International Conference on Advanced Learning Technologies, Timisoara, Romania. https://doi.org/10.1109/ICALT.2017.40

Aronson, L. (2011). Twelve tips for teaching reflection at all levels of medical education. *Medical Teacher, 33*(3), 200–205. https://doi.org/10.3109/0142159X.2010.507714

Atwood, S. A., Siniawski, M. T., & Carberry, A. R. (2014). *Using standards-based grading to effectively assess project-based design courses.* Paper presented at the ASEE Annual Conference and Exposition, Indianapolis, IN. https://doi.org/10.18260/1-2–23278

Badir, A., & O'Neill, R. (2017). *Homework graded by students.* Paper presented at the ASEE Annual Conference and Exposition, Columbus, OH. https://doi.org/10.18260/1-2–28440

Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*(2), 213–238. https://doi.org/10.3102/00346543061002213

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Boud, D. (2017). Standards-based assessment for an era of increasing transparency. In D. Carless, S. Bridges, C. K. Y. Chan, & R. Glofcheski (Eds.), *Scaling up assessment for learning in higher education: The enabling power of assessment* (Vol. 5, pp. 3–17). Springer. https://doi.org/10.1007/978-981-10-3045-1

Boyd, E. M., & Fales, A. W. (1983). Reflective learning: Key to learning from experience. *Journal of Humanistic Psychology, 23*(2), 99–117. https://doi.org/10.1177/0022167883232011

Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education, 36*(4), 35–407. https://doi.org/10.1080/03075071003642449

Chickering, A. W., & Ehrmann, S. C. (1996). Implementing the seven principles: Technology as lever. *AAHE Bulletin, 49*(2), 3–6.

Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin, 39*, 3–7. https://doi.org/10.1016/0307-4412(89)90094-0

Chizmar, J. F., & Ostrosky, A. L. (1998). The one-minute paper: Some empirical findings. *The Journal of Economic Education, 29*(1), 3–10. https://doi.org/10.1080/00220489809596436

Diefes-Dux, H. A. (2016). *Student reflections on standards-based graded assignments.* Paper presented at the Frontiers in Education Conference, Erie, PA. https://doi.org/10.1109/FIE.2016.7757445

Diefes-Dux, H. A. (2019). Student self-reported use of standards-based grading resources and feedback. *European Journal of Engineering Education, 44*(6), 838–849. https://doi.org/10.1080/03043797.2018.1483896

Diefes-Dux, H. A., & Carberry, A. R. (2019a). *Student reflections on proficiency with learning objectives: Early semester actions and plans.* Paper presented at the Research in Engineering Education Symposium, Cape Town, South Africa. Retrieved from https://www.reen.co/papers-proceedings

Diefes-Dux, H. A., & Carberry, A. R. (2019b). *Cases of student reflection within a course using standards-based grading.* Paper presented at the Frontiers in Education Conference, Cincinnati, OH. https://doi.org/10.1109/FIE43999.2019.9028501

Diefes-Dux, H. A., & Cruz Castro, L. M. (2018). *Student reflection to improve access to standards-based grading feedback.* Paper presented at the Frontiers in Education Conference, San Jose, CA. https://doi.org/10.1109/FIE.2018.8659325

Diefes-Dux, H. A., & Cruz Castro, L. M. (2019a). *Patterns of monthly student access to feedback by section in a large course using standards-based grading and reflection.* Paper presented at the Research in Engineering Education Symposium, Cape Town, South Africa. Retrieved from https://www.reen.co/papers-proceedings

Diefes-Dux, H. A., & Cruz Castro, L. M. (2019b). *Validation of an instrument to measure student engagement with a standards-based grading system.* Paper presented at the ASEE Annual Conference and Exposition, Tampa, FL. https://doi.org/10.18260/1-2–33530

Diefes-Dux, H. A., & Ebrahiminejad, H. (2018). *Standards-based grading derived data to monitor grading and student learning.* Paper presented at the ASEE Annual Conference and Exposition, Salt Lake City, UT. https://doi.org/10.18260/1-2–30981

Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research, 83*(1), 70–120. https://doi.org/10.3102/0034654312474350

Faraway, J. J. (2016). *Extending the linear model with R: Generalized linear, mixed effects and nonparametric regression models* (2nd ed.). CRC Press. https://doi.org/10.1201/9781315382722

Gibbs, G., & Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education, 2004-05*(1), 3–31. Retrieved from http://eprints.glos.ac.uk/3609/

Guskey, T. R. (2001). Helping standards make the grade. *Educational Leadership, 59*(1), 20–27.

Handley, K., Price, M., & Millar, J. (2011). Beyond 'doing time': Investigating the concept of student engagement with feedback. *Oxford Review of Education, 37*(4), 543–560. https://doi.org/10.1080/03054985.2011.604951

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Heywood, J. (2014). *The evolution of a criterion referenced system of grading for engineering science coursework.* Paper presented at the Frontiers in Education Conference, Madrid, Spain. https://doi.org/10.1109/FIE.2014.7044236

Hicks, N. M. (2020). *A variability analysis of grading open-ended tasks with rubrics across many graders* [Doctoral dissertation, Purdue University]. ProQuest Dissertations Publishing.

Hicks, N. M., & Douglas, K. A. (2018). *Efforts to improve undergraduate grader consistency: A qualitative analysis*. Paper presented at the ASEE Annual Conference and Exposition, Salt Lake City, UT. https://doi.org/10.18260/1-2–30366

Hounsell, D., McCune, V., Hounsell, J., & Litjens, J. (2008). The quality of guidance and feedback to students. *Higher Education Research & Development, 27*(1), 55–67. https://doi.org/10.1080/07294360701658765

Klobas, J. E., & McGill, T. J. (2010). The role of involvement in learning management system success. *Journal of Computing in Higher Education, 22*(2), 114–134. https://doi.org/10.1007/s12528-010-9032-5

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284. https://doi.org/10.1037/0033-2909.119.2.254

Laflen, A., & Smith, M. (2017). Responding to student writing online: Tracking student interactions with instructor feedback in a learning management system. *Assessing Writing, 31*, 39–52. https://doi.org/10.1016/j.asw.2016.07.003

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics, 34*(1), 1–14. https://doi.org/10.2307/1269547

Lee, E., Carberry, A., Diefes-Dux, H. A., & Atwood, S. (2018). Faculty perception before, during, and after implementation of standards-based grading. *Australasian Journal of Engineering Education, 23*(2), 53–61. https://doi.org/10.1080/22054952.2018.1544685

Lizzio, A., & Wilson, K. (2008). Feedback on assessment: students' perceptions of quality and effectiveness. *Assessment & Evaluation in Higher Education, 33*(3), 263–275. https://doi.org/10.1080/02602930701292548

Marzano, R. J. (2010). *Formative assessment & standards-based grading*. Marzano Resources.

Mendez, J. (2018). *Standards-based specifications grading in a hybrid course*. Paper presented at the ASEE Annual Conference and Exposition, Salt Lake City, UT. https://doi.org/10.18260/1-2–30982

Mensink, P. J., & King, K. (2020). Student access of online feedback is modified by the availability of assessment marks, gender and academic performance. *British Journal of Educational Technology, 51*(1), 10–22. https://doi.org/10.1111/bjet.12752

Montero, P., & Vilar, J. A. (2014). TSclust: An R package for time series clustering. *Journal of Statistical Software, 62*(1), 1–43. https://doi.org/10.18637/jss.v062.i01

Mulliner, E., & Tucker, M. (2017). Feedback on feedback practice: Perceptions of students and academics. *Assessment & Evaluation in Higher Education, 42*(2), 266–288. https://doi.org/10.1080/02602938.2015.1103365

Mutch, A. (2003). Exploring the practice of feedback to students. *Active Learning in Higher Education, 4*(1), 24–38. https://doi.org/10.1177/1469787403004001003

Nicol, D. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. *Assessment & Evaluation in Higher Education, 35*(5), 501–517. https://doi.org/10.1080/02602931003786559

Nilson, L. B. (2014). *Specifications grading: Restoring rigor, motivating students, and saving faculty time.* Stylus Publishing.

Post, S. L. (2014). *Standards-based grading in a fluid mechanics course.* Paper presented at the ASEE Annual Conference and Exposition, Indianapolis, IN. https://doi.org/10.18260/1-2–23032

Price, M., Handley, K., Millar, J., & O'Donovan, B. (2010). Feedback: All that effort, but what is the effect? *Assessment & Evaluation in Higher Education, 35*(3), 277–289. https://doi.org/10.1080/02602930903541007

Quinton, S., & Smallbone, T. (2010). Feeding forward: Using feedback to promote student reflection and learning—A teaching model. *Innovations in Education and Teaching International, 47*(1), 125–135. https://doi.org/10.1080/14703290903525911

Ridout, M., Hinde, J., & Demetrio, C. G. (2001). A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. *Biometrics, 57*(1), 219–223. https://doi.org/10.1111/j.0006-341X.2001.00219.x

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119–144. Retrieved from http://www.jstor.org/stable/23369143

Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education, 30*(2), 175–194. https://doi.org/10.1080/0260293042000264262

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action.* Basic Books.

Scriffiny, P. L. (2008). Seven reasons for standards-based grading. *Educational Leadership, 66*(2), 70–74.

Selbach-Allen, M. E., Greenwald, S. J., Ksir, A. E., & Thomley, J. E. (2020). Raising the bar with standards-based grading. *Problems, Resources, and Issues in Mathematics Undergraduate Studies* (*PRIMUS*), 30(8–10), 1110–1126. https://doi.org/10.1080/10511970.2019.1695237

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Stratman, E., & Diefes-Dux, H. A. (2021). *Impact of weekly reflection prompts on first-year engineering students' metacognitive strategies* [Poster presentation]. Nebraska Virtual Summer Research Symposium 2021, Lincoln, NE. Retrieved from https://mediahub.unl.edu/media/17364

The R Foundation. (n.d.). *Getting started.* The R Project for Statistical Computing. Retrieved from https://www.r-project.org/

Vuong, Q. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, *57*(2), 307–333. https://doi.org/10.2307/1912557

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis.* Springer-Verlag. https://doi.org/10.1007/978-3-319-24277-4

Wierer, J. (2019). *Standards-based grading for signals and systems.* Paper presented at the ASEE Annual Conference and Exposition, Tampa, FL. https://doi.org/10.18260/1-2–33282

Wiggins, G. (2012). Seven keys to effective feedback. *Educational Leadership*, *70*(1), 10–16.

Winter, C., & Dye, V. L. (2004). An investigation into the reasons why students do not collect marked assignments and the accompanying feedback. In H. Gale (Ed.), *Learning and teaching projects 2003/04* (pp. 133–141). University of Wolverhampton Press.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, *25*(1), 3–17. https://doi.org/10.1207/s15326985ep2501_2

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Author Biographies

Heidi A. Diefes-Dux is a professor of engineering education in the Department of Biological Systems Engineering at the University of Nebraska–Lincoln, East Campus, 210 L. W. Chase Hall, Lincoln, NE 68583-0726, USA; heidi.diefes-dux@unl.edu

Laura M. Cruz Castro is a doctoral student in the School of Engineering Education at Purdue University, 3577 Seng-Liang Wang Hall, 516 Northwestern Avenue, West Lafayette, IN 47906, USA; lcruzcas@purdue.edu