Optimising Automated Feedback Systems to Motivate Students

Peter Mansfield
Edith Cowan University, Perth, Australia
p.mansfield@ecu.edu.au

Daniel Boase-Jelinek
Murdoch University, Perth, Australia
d.boase-jelinek@murdoch.edu.au

We have developed a customisable, web-based proficiency test system that delivers randomised questions, contextual data, and formative feedback to students as part of the Business Edge program for Bachelor of Business students at Edith Cowan University. Our primary aim in developing this system was to create a web-based learning environment that promotes motivation and learning through controlled formative feedback.

Additional aims included: the development of easy-to-use lecturer administration; automated record keeping; automated marking; and the ability to offer students multiple attempts on a given test (with new questions / datasets for each attempt). These 'lower-level aims' were met successfully.

Nevertheless, there is more to be done: we did not reach the theoretically ideal learning environment required to achieve our desired outcomes to promote motivation and learning based upon formative feedback.

We conclude that the formative feedback components of our proficiency test system need to be enhanced significantly if it is to be of true benefit to students.

1 Introduction

Giving students formative feedback is potentially the single most effective intervention that a teacher can make for enhancing student learning (Hattie, 1999). We ask: how can teachers best use formative feedback to help their students reach their potential? How does it affect their motivation to learn? What happens when the ideal learning environment for giving formative feedback and motivating students is not achieved within an automated feedback system? In this paper we outline our efforts to find answers to some of these questions as we move towards our goal of further development of the automated formative feedback system we have developed.

1.1. Defining formative feedback

Feedback may be defined as a mechanism that machines and computer systems use to maintain or reach a desired state (Lewis, 1992). The feedback mechanism operates by measuring the current state of a system and comparing it to the desired state of the system; the difference between these two measurements being used to adjust performance to bring the system closer to the desired state.
Control systems theory may be applied to human learning as well as machines (Abulwahed, Nagy, & Blanchard, 2009). For example, a constructivist student-centred learning environment may be characterised as a closed-loop learning system to the extent that the learning environment provides feedback to the student about the gap between actual and desired performance and how to close that gap. Using models of a constructivist student-centred learning environment, Abulwahed et. al. (2009) found that a feedback loop can help weak students reach the same learning outcomes as

strong students. In contrast, weak students without formative feedback did not achieve their learning outcomes within the desired time frame (if at all).
This finding is supported by a meta-study of educational interventions (Black & Wiliam, 1998) which found that improved formative assessment helps weak learners more than it helps strong learners.
Formative assessment may be defined as a 'control system' having three elements: "the desired goal, the evidence about their present position, and some understanding of a way to close the gap between the two" (Sadler, 1989). Formative feedback is the information component of the formative assessment system whereby the system is informed about (a) the discrepancy between a desired and actual state, and (b) the changes required to the system to bring it closer to the desired state.
From control systems theory we therefore have a view of formative feedback as being the information component of a formative assessment-based learning environment that sets clear, observable goals for students, informs them about how close they are to reaching these goals, and what they need to change in order to do so.

1.2. Providing effective formative feedback

Effective formative feedback must be given in the context of a formative assessment-based learning environment that has a number of characteristics (Bangert-Drowns, C. C. Kulik, J. A. Kulik, & Morgan, 1991; Black & Wiliam, 1998; Hattie, 1999). These characteristics are most likely to be met when teachers:

- create regular challenges for students,
- set criteria for successful completion of those challenges,
- provide criteria that are specific, challenging (within each students' capability, but beyond current level of performance), and measurable,
- give feedback based on the changes required to achieve the desired level of performance,
- delay giving feedback until the student has exhausted all avenues for addressing the challenge,
- encourage students to self-assess themselves before being given feedback,
- use student self-assessment as a basis for classroom discussion, and
- provide summative assessment as an objective measure of performance.

In contrast, formative feedback can inhibit learning in certain situations (Bangert-Drowns et al., 1991). These include giving feedback where:

- students have not yet exhausted attempts to address the challenge,
- students are not confident about their response to the challenge, or
- the challenge is too easy or too difficult.

As class sizes grow so does the difficulty of setting up a formative assessment process providing timely feedback that focuses on the individual changes each student needs to make to address a challenge. It is not surprising then that automated approaches to providing formative feedback are being explored (Barker, 2010; Blayney & Freeman, 2004; Hatziapostolou & Paraskikis, 2010; Sondergaard & Thomas, 2004).

1.3. Automating formative feedback

Automated feedback systems require information about what a student is doing, and what a student should be doing. For example, (Blayney & Freeman, 2004) describe an Excel-based formative feedback system in which students enter Excel formulas for solving a problem and the system tells them whether or not their formula is correct. If their answer is not correct it attempts to suggest the kind of error made, and allows students to make another attempt. Even in more complex calculations such as open-

ended computer programming challenges it is possible to develop algorithms for evaluating the quality of student work and giving them automated feedback reliably (Boonprasert, 2009).
The major challenge of automated formative feedback systems is to give accurate feedback. In one study (Blayney & Freeman, 2004), only just over half (55%) of students felt that the feedback accurately assessed their attempts, but this figure was confounded by an error in the marking algorithm that was later corrected. Whilst these problems are avoidable, it seems that student confidence in the system is sensitive to any problems that arise.
Automated feedback systems can motivate students to seek to reach the specified goals. In the Blayney & Freeman (2004) study, students attempted each assignment an average of 10 times. Some students used the system many more times (one student made 97 attempts during one week).
Such systems can only give feedback about how to improve the solution if students are given a space in which to do their working-out as well as entering their solution to the challenge. It becomes much more difficult to provide effective feedback in situations where only the students' answers are provided and none of their solution method.

1.4. Formative feedback and motivation

Whilst control systems theory may provide a useful guide for structuring formative feedback to help students reach their goals, it says little about their motivation to do so. One theoretical perspective that may prove useful in this context is students' goal orientation (Beghetto, 2004). Goal orientations refer to the students' reasons for engaging in an activity. The two types of goal orientations typically identified are: *mastery goals* and *performance goals*.
Students with an orientation towards *mastery goals* generally focus on understanding a problem and seeking a solution to it (Beghetto, 2004). Such students tend to view formative feedback as an opportunity to learn, and are more likely to persevere until they have achieved mastery.
Students with a *performance goal* orientation are likely to either take low-risk strategies (*performance avoidance*) such as cheating to avoid failure, or high-risk strategies (*performance approach*) by competing for the highest scores (Elliot & Moller, 2003). The *performance approach* orientation is motivational only while it works (i.e. the student is able to get good marks). Students with a *performance goal* orientation tend to perceive formative feedback as criticism and a sign of failure. It seems therefore that the key to motivating poor performing students to do better is to foster a *mastery goal* orientation (Beghetto, 2004).
A *mastery goal* orientation can be fostered with appropriate classroom conditions (Ben-Ari & Eliassy, 2003) that involve:

- setting challenging tasks with clear goals,
- evaluating students on the basis of effort, progress and mastery,
- developing a culture that perceives mistakes as a legitimate and necessary part of learning, and
- encouraging students to monitor their own progress

1.5. Applying theory to the Proficiency Tests in BES1200

It seems that the kind of learning environment required to provide effective formative feedback has much in common with the kind of learning environment that motivates students to persevere until they have mastered the challenge presented to them. It is not so clear whether automated formative feedback systems can provide the kind of learning environment required to foster a *mastery goal* orientation. It is also not clear what will happen in situations (such as the Proficiency Tests in BES1200) where students only enter their solutions, and not their solution method.

2 The BES1200 Context

The Faculty of Business and Law of Edith Cowan University launched its new Business Edge Program in 2007. The key purpose of this program is to develop the employability skills of all students pursuing a Bachelor of Business degree. ECU views these employability skills as generic rather than discipline-specific; the Business Edge program resides within the Faculty's Dean's Office, rather than within one of the Faculty's four Schools.

The program is comprised of four units; it is expected that students will complete the first two units during their first year of study. The third and fourth units are usually completed during students' second and third years of study, respectively.

2.1. Overview of Unit

All ECU Business Edge Units are designed to help students develop skills desired in the workplace. The second of the four units – BES1200 Business Knowledge Development – is usually taken by first year students during their second semester of study.

The number of students taking this unit in a given semester often exceeds four hundred. Students meet once a week with a facilitating lecturer for three hours. The typical class size is approximately 25.

The key deliverable for the unit BES1200 is a 'Final Team Report'. This report reflects the efforts of teams of three to five students to conceive and develop a new business.

As the semester unfolds and teams develop their business ideas, students are assessed individually on: the initiative they demonstrate; their ability to assess the businesses of other teams; their communication skills, both written and oral; and their ability to analyse simplified datasets extracted from ABS census data that is augmented by hypothetical market survey data.

2.2. Data analysis

In BES1200 the ability to analyse data is developed through a series of Case Studies that highlight basic statistical concepts. Detailed information about the Case Studies and the associated Proficiency Tests is provided in the next Section.

We note that the proposed businesses and the developed statistical skills are linked: as part of the business development project, teams design and implement a market survey of prospective customers. Each team collects data face to face. It then analyses the data it collects to assess potential demand for its products and/or services.

2.3. Case Studies and the Proficiency Tests

Utilising our generalised testing system, we have developed a customised series of Proficiency Tests (PTs) for BES1200. Student performance on these tests constitutes 30% of the total assessment for the Unit.

These tests are based upon a series of Case Studies that we have developed in-house. We discuss these Case Studies to provide context, then follow with a discussion of several features of the Proficiency Tests.

2.4. The Case Studies that are part of the Business Edge Unit BES1200

The Case Studies (CSs) are Excel-based. They cover the basics of statistical analysis. They have been developed in-house. Students undertake them in class, using ECU-provided laptops. They are self-paced; an hour of class time is allocated for each of them. Discussion amongst students is encouraged.

The datasets analysed in the CSs are a combination of simplified census data and hypothetical responses to market surveys. There are two main desired learning outcomes sought through completion of the CSs; the first is to provide students with the skills necessary to analyse the actual survey data they collect face to face in the course of developing their proposed businesses. The second is to develop students' Excel skills.

The following table summarises the material covered in each of the CSs. It is expected that successful completion of a given CS prepares students to undertake the associated Proficiency Test (last column). Details about the PTs follow the table.

Table 2-1: Case Studies and the Proficiency Tests

Case Study (CS)	Material Covered (Excel based)	Proficiency Test (% of assessment; how to access; access time)
CS0	Tabular data; sorting data; median; average	PT0 (practice)
CS1	Basic Excel functions: arithmetic operations, Excel's SUM function	PT1 (5%; online; one week)
CS2	Average; Median; Standard Deviation	PT2 (5%; online; one week)
CS3A	Covariance; Correlation	PT3 (5%; online; one week)
CS3B	Linear best fit; regression	
-	All of the above material	PT4 (15%; in-class; 90 minutes)

3 The Proficiency Tests

3.1. Formative feedback

We have developed a customisable web-based proficiency test system that delivers randomised questions, contextual data, and formative feedback to students. The formative feedback provided to students can be adjusted.

The level of formative feedback begins with: no feedback. At the next level, only the number of questions correctly answered is provided. This level is followed by information about which questions were answered correctly; at the next level, the correct answers to each question are provided.

The most detailed level of formative feedback provides students with the correct solution method to each problem.

Our system includes controls that enable immediate feedback upon test completion as well as delayed feedback at the end of a test period.

Our system does not collect information about the formulas students used to find an answer to each problem, and is therefore unable to give feedback to students about their solution method. We note here that the implementation of such feedback is not likely to be straightforward.

3.2. Context and operations

Each of the Proficiency Tests (PTs) is available following the associated Case Study (see Table 2-1: Case Studies and the Proficiency Tests). Each class of students is provided with a password to help streamline lecturers' administration.

Students download a new Excel spreadsheet of data each time they attempt a given PT; questions are selected at random from a large pool of questions. The associated datasets also change.

The web-based PT questions require analysis of the dataset contained on the downloaded Excel spreadsheet. Students analyse the data on their desktops using features of Excel covered in the associated CS and then enter their answers onto the PT web page.

The first PT (PT0) is not assessable. It is undertaken in class, using ECU-provided laptops. The main purpose of the test is to provide students with the opportunity to step through the technology in preparation for the PTs that follow later in the semester.

The next three PTs are undertaken online by students outside of class. Each of them is available for one week following the discussion of the associated Case Study.

The questions on each instance of each PT range in difficulty, from easy through challenging. Point values for each correct answer are allocated in advance, with the more difficult questions attracting more points.

At the outset of this program we considered how to set the controls of our system to motivate students to do their best. After some consideration we elected to adopt the following settings. Students were informed in advance of these settings.

3.3. Settings used for online tests (Proficiency Tests 1-3)

The feedback settings used were:

(a) Immediate feedback:
Students were given the opportunity to attempt each of these three online tests multiple times; for each attempt, the questions and the associated dataset changed. Upon completion of each attempt, the student received immediate notification of his/her grade. The student was not informed which questions were answered correctly and which were answered incorrectly.

(b) Delayed feedback:
At the end of the time period for test access, students are provided with detailed feedback about every question on every test. Short videos were made available to students to explain how to do each question.

3.4. Rationale for settings for online tests

It was our belief that students would welcome the opportunity to take a given test multiple times, with only their best results counting towards assessment ... and that by learning their overall scores immediately, they would make new attempts until they were satisfied with their best scores. We knew that some of them might find it frustrating that they were not informed at the time about which questions they answered correctly / incorrectly ... but we thought that this lack of detail would prod them to reflect carefully on their answers as they made new attempts.

Students were allowed 100 attempts at each of the three online tests. They were given one week to complete each set of attempts. The tests were untimed.

3.5. Settings used for in-class test (Proficiency Test 4)

PT4 was a timed test (90 minutes). Only one attempt was permitted, and no feedback was provided at the completion of the test.

4 Students' test performances and perseverance

We outlined in Section 3.4 our rationale for the test control and formative feedback settings we used in this research: students had the opportunity to take each of the three online tests as often as they would like, with only the best result counting towards their final assessments; each time they undertook a given Proficiency Test, they were informed immediately of their overall result only.

In this section we summarise the relationships we have uncovered between number of attempts (we take this number as a measure of *observed perseverance*) and actual test performance. As a prelude, we first discuss the main reason that our analysis has some unavoidable limitations.

4.1. Prelude: a caveat

In brief, it is likely that there are many students who are perfectly satisfied with their best result on a given Proficiency Test after a small number of attempts. Indeed, it is certainly possible that those students who are able and prepared to undertake a given PT attain a very high mark on their first or second attempt; such students would have no reason to make further attempts. To the extent that the number of such students is significant, their inclusion in the data necessarily dilutes any quantitatively based argument we might be able to make that student performance goes up with increasing number of attempts.

We move on to our analysis of the data.

4.2. Test performance and perseverance for PT1

It was certainly our expectation at the outset of this research that there would be a positive correlation between the number of attempts made and best score.

Figure 4-4.1 displays the scatter plot of number of attempts made on PT1 versus best score attained. Each marker on the figure represents one student. Note that the horizontal axis is logarithmic.

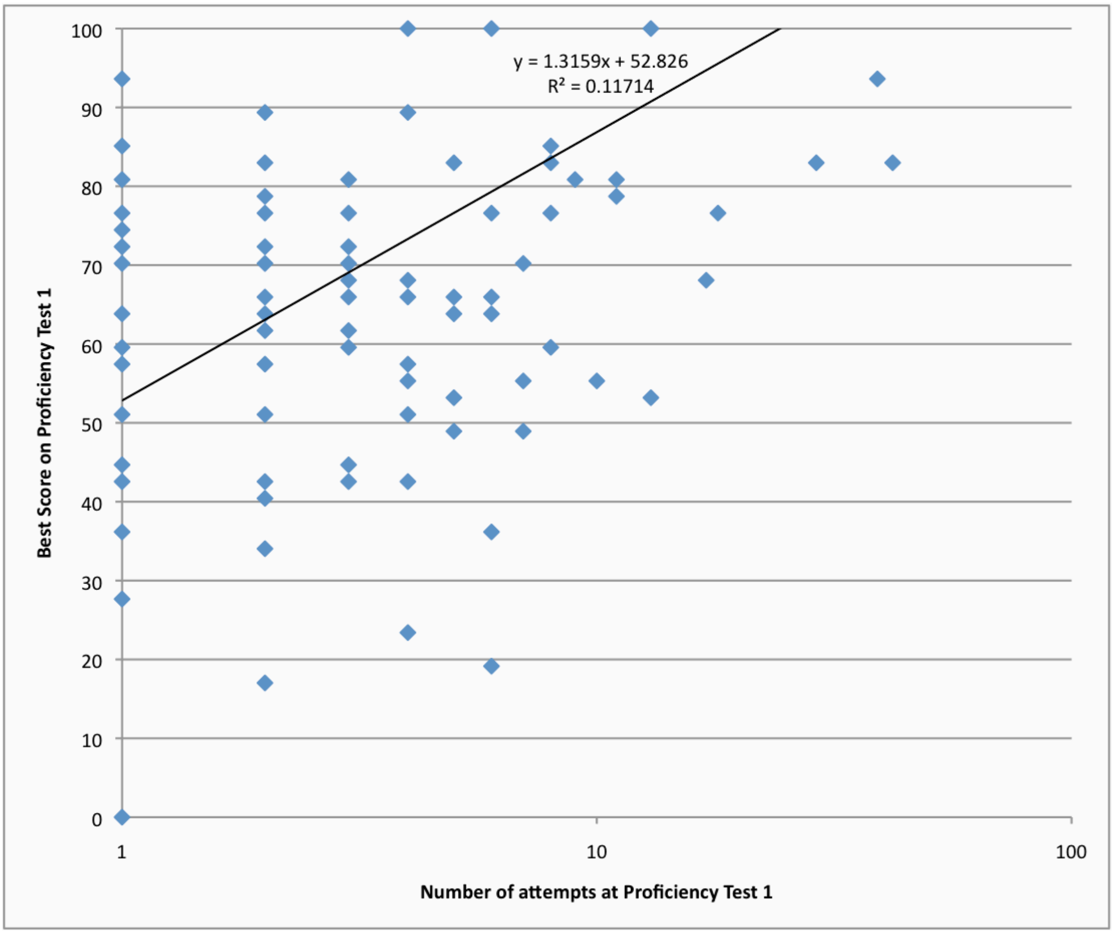

Figure 4-4.1 : Correlation between number of attempts and test score on proficiency test 1.

Superimposed on the scatter plot is the regression line together with its equation. The main feature of the regression line is its positive slope[i]; see Table 4-1.

Table 4-1: Regression line parameters for Figure 4-4.1

	Slope	Intercept
coefficient	1.32	52.83
standard error	0.36	2.07

Similar results hold for the second and third Proficiency Tests.

Despite the caveat we presented above in Section 4.1, we have established a positive relationship between number of attempts (observed perseverance) and best score. This relationship would be strengthened further if there was a way to exclude from the data those relatively able students who might have been satisfied with their scores after a small number of attempts.

In the next section we discuss some of the main results of our survey of students' attitudes.

5 Survey results

5.1. Overview

We thought it would be useful to survey students so that we could develop an understanding of their attitudes towards the content and skills covered in the Case Studies and the associated Proficiency Tests.

Our first survey (implemented at the time the Case Studies began in class) explored a range of issues related to statistics and Excel. Students were also asked to provide self-assessments of their ability to persevere.

We implemented a second survey at the end of the last Proficiency Test. Although the responses to this survey are not analysed in this paper, we note here that of the nearly 200 students who were enrolled in the Unit during the semester, 104 of them completed both survey instruments and undertook the majority of the Proficiency Tests. The results we report in the following sections are based upon the responses of these 104 students[ii].

For the purposes of this paper, we focus on students' self-assessments of their ability to persevere. The analysis of their responses about statistics and Excel will be addressed in a separate paper.

5.2. Students' self-assessments of their ability to persevere

In this section we address students' self-assessment of their ability to persevere. Here is the key multiple choice question they were asked to answer, followed by a chart of their answers (the text of the four possible answers are included along the horizontal axis of the figure following (Figure 5.1)).

Question: Please rate your ability to persevere at your studies, e.g., to try and try again:

Responses:

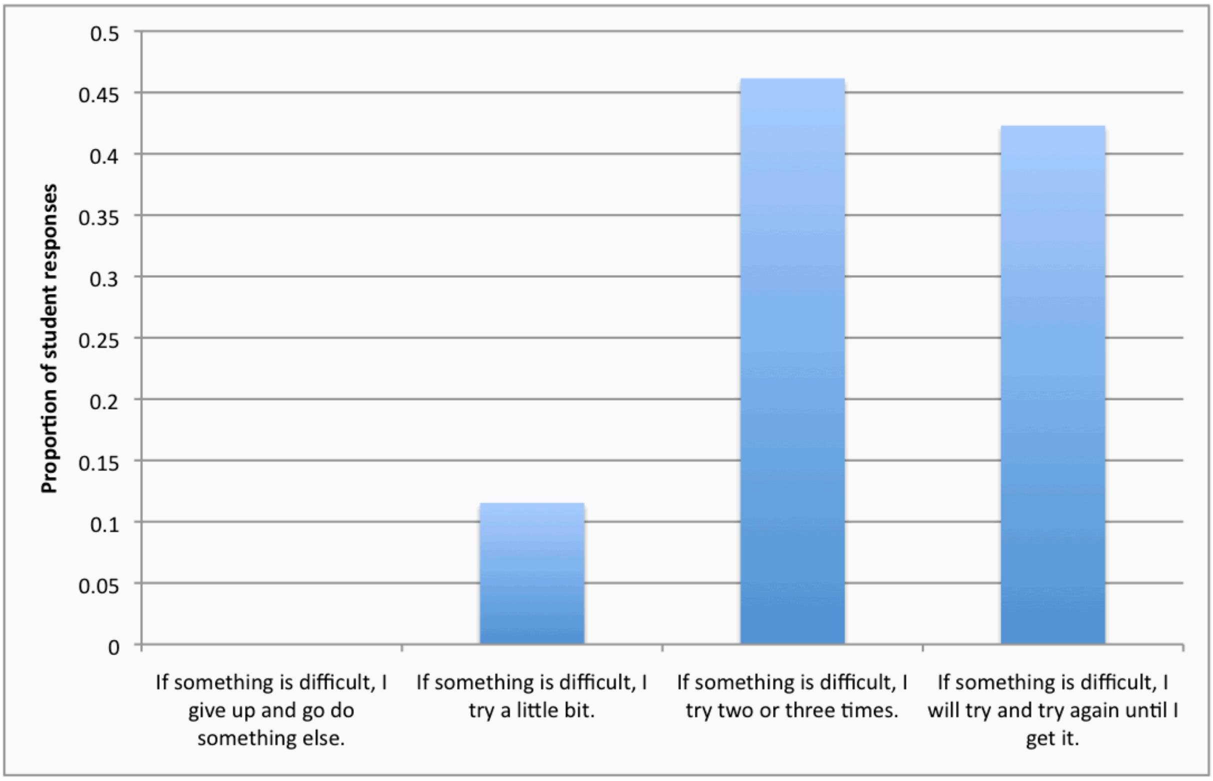

Figure 5.1: Claimed Perseverance

We are interested in linking performance with self-assessment of perseverance; we do so in the next section.

5.3. Linking performance with self-assessment of perseverance

It is very interesting to consider the question: "To what extent does self-assessment of perseverance correlate with actual effort?"

It is not possible for us to answer this question with complete satisfaction – we addressed this issue in Section 4, where we commented that it is not possible to measure students' actual ability to persevere if they attain high marks on a given PT after one or two attempts. Nevertheless, we can make some inroads.

We make a few comments to help the reader understand the three tables that follow.

We have segmented our respondents into two groups: the first group is comprised of those students who self-assess their perseverance as *very high* ("If something is difficult, I will try and try again until I get it."); the second group is comprised of those students who self-assess their perseverance as lower than this (*less than very high*). There are 44 students (42.3% of all respondents) in the first group and 60 students (57.7% of all respondents) in the second.

Table 5-1: Self-assessed perseverance and number of attempts for Proficiency Test 1

Self-assessment of perseverance	Average number of attempts	Average score
Very high	6.4	57.7
Less than very high	3.3	59.2

Table 5-2: Self-assessed perseverance and number of attempts for Proficiency Test 2

Self-assessment of perseverance	Average number of attempts	Average score
Very high	3.7	68.2
Less than very high	3.0	68.4

Table 5-3: Self-assessed perseverance and number of attempts for Proficiency Test 3

Self-assessment of perseverance	Average number of attempts	Average score
Very high	2.6	61.0
Less than very high	1.6	48.5

There are a number of features about the findings summarised in these three tables that warrant comment.

- There is a positive correlation between self-assessment of performance and the average number of attempts made on each Proficiency Tests;
- The number of attempts decreases as the semester unfolds (from PT1 to PT2 to PT3)[iii];
- For the first two PTs, there is no apparent relationship between self-assessment of perseverance or average number of attempts and score attained;
- For the third PT, there is (finally) evidence that perseverance pays off; the average score for the group that self-assesses very high perseverance is significantly higher than the average score for the other group.

At the time we constructed our survey instrument we had the unstated expectation that 'perseverance' would apply to a single task (i.e., to a single PT). The evidence supports the idea that this notion of perseverance is positively correlated to outcome. Yet the evidence seems to suggest that there is an additional element of perseverance: students in Group 1 who self-assess very high perseverance ultimately 'triumph' on PT3, as evidenced by their significantly higher average score for this test.

And yet ... this triumph of perseverance appears to be short-lived.

Consider the following table for the final, in-class Proficiency Test (PT4) (only one attempt allowed): we see that there is no real difference in average score achieved for the two groups. In other words: there is no evidence based on the results of this test that self-assessed highly persevering students who made more attempts on the online tests have mastered the underlying material any better than the other students.

Table 5-4: Self-assessed perseverance and number of attempts for Proficiency Test 4

Self-assessment of perseverance	Average number of attempts	Average score
Very high	1	75.4
Less than very high	1	74.4

We can offer no explanation why perseverance does not appear to have 'paid off' for those students in our first group, other than one possibility: perhaps PT4 was too easy (the average scores are very high as compared to the average scores for the earlier tests).

6 Discussion

We concede that the focus of our automated formative feedback system set up to respond to student answers rather than the method used to derive the answers is not consistent with the kinds of learning environments recommended learning by either control systems theory (as applied to learning), or goal orientation theory. This divergence from the recommendations of theory may provide part of the explanation for some of the observations we have made from the data:

- Despite the positive correlation between number of attempts at a test and test scores, the number of attempts at proficiency tests declined during semester (Table 5-1, Table 5-1, and Table 5-2).
- The average number of attempts (for high persevering students) dropped from 6.4 in PT1 to 3.7 for PT2 and 2.6 for PT3. Similarly, the number of attempts (for low persevering students) fell from 3.3 for PT1 to 3.0 for PT2 and 1.6 for PT3.

Overall, it appears that we did not succeed in motivating students to try harder (make more attempts) at the proficiency tests. This occurred despite the fact that those students who did make more attempts were more likely to do better at the test.
The apparent decline in number of attempts by (both high and low persevering) students is consistent with the predictions of goal orientation theory that a focus on scores rather than method will foster a *performance goal* orientation amongst students, and result in poor motivation in the case of lack of performance. It is even possible that low persevering students stayed (or became even more) *performance goal* oriented, and thus lost motivation during semester. This is the group of students who should benefit most in motivation when the classroom environment fosters a *mastery goal* orientation.

7 Conclusions and future directions

It seems that when automated formative feedback systems focus on correct answers rather than the solution method adopted by students then it fosters a *performance goal* orientation amongst the students that may be counter-productive to their motivation and learning outcomes. Our challenge in developing the proficiency test system is to find a way to respond to these findings by developing ways to:

- clearly specify criteria for mastery;
- provide space for students to specify how they went about solving a problem;
- develop algorithms for assessing the problem solutions methods used; and
- giving meaningful feedback to students about those methods.

8 References

Abulwahed, M., Nagy, Z., & Blanchard, R. (2009). The Feedback Impact on Learning, A Control Systems View. Presented at the 20th Australian Association for Engineering education Conference, Adelaide: Australian Association of Engineeering Education.

Bangert-Drowns, R., Kulik, C. C., Kulik, J. A., & Morgan, M. T. (1991). The Instructional Effect of Feedback in Test-like Events. *Review of Educational Research*, *61*(2), 213-238.

Barker, T. (2010). An Automated Feedback System Based on Adaptive Testing: Extending the Model. *International Journal of Emerging Technologies in Learning*, *5*(2), 11-14.

Beghetto, R. A. (2004). Toward a more complete picture of student learning: assessing students' motivational beliefs. *Practical Assessment, Research & Evaluation*, *9*(15). Retrieved from http://pareonline.net/getvn.asp?v=9&n=15

Ben-Ari, R., & Eliassy, L. (2003). The differential effects of the learning environment on student achievement motivation: A comparision between frontal and complex instruction strategies. *Social Behaviour and Personality*. Retrieved from http://findarticles.com/p/articles/mi_qa3852/is_200301/ai_n9223947/?tag=content;col1

Black, P., & Wiliam, D. (1998). Inside the Black Box: Raising standards through classroom assessment. *Phi Delta Kappan*, *80*(2), 139-148.

Blayney, P., & Freeman, M. (2004). Automated formative feedback and summative assessment using individualised spreadsheet assignments. *Australasian Journal of Educational Technology*, *20*(2), 209-231.

Boonprasert, S. (2009). The Development of Automatic Marking System for Formative Assessment. Presented at the 13th UNESCO-APEID International Conference on Education and World Bank-KERIS High Level Seminar on ICT in Education, Hangzhou, People's Rebuplic of China: UNESCO. Retrieved from http://www.unescobkk.org/fileadmin/user_upload/apeid/Conference/13th_Conference/Papers/1.E.3_Automatic_Marking_System_for_Formative_Assessment__Siwanan_.pdf

Elliot, A. J., & Moller, A. C. (2003). Performance-approach goals: good or bad forms of regulation? *International Journal of Educational Research*, *39*(4-5), 339-356.

Hattie, J. (1999, August 2). Influences on student learning. Retrieved October 14, 2010, from http://www.geoffpetty.com/downloads/WORD/Influencesonstudent2C683.pdf

Hatziapostolou, T., & Paraskikis, I. (2010). Enhancing the Impact of Formative Feedback on Student Learning Through an Online Feedback System. *Electronic Journal of eLearning*, *8*(2), 51-90.

Lewis, F. L. (1992). A brief history of feedback control - Chapter 1. Retrieved October 14, 2010, from http://www.theorem.net/theorem/lewis1.html

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119-144. doi:10.1007/BF00117714

Sondergaard, H., & Thomas, D. (2004). Effective Feedback to Small and Large Classes. Presented at the 34th ASEE/IEEE Frontiers in Education Conference, Savannah, GA, USA: IEEE. Retrieved from ieeexplore.ieee.org/iel5/9652/30543/01408573.pdf

---

[i] The t-value for the slope coefficient is 3.678; this corresponds to a p-value less than 0.0005.

[ii] We made the decision to restrict our analysis in this paper to the responses and the performances of the 104 students who undertook both surveys for a reason: by doing so, our analysis of any future discovered results based on both surveys will be completely consistent with all of the results in this paper.

[iii] This drop-off can be partially attributed to the fact that the first Proficiency Test is considered by students to be the most difficult one of the three (we concur).