The Effects of Automated Email-Feedback on Final Exam Grades and Dropout Rates

Alline Aleman

S3132900

Master Thesis Learning in Interaction PAMAOW05

Faculty of Behavioural and Social Sciences, University of Groningen

Supervisor: N. Renting

Second Assessor: J. Brouwer

Submitted: June 22, 2023

Word Count: 5482

Abstract

Educational scientists agree that feedback is a crucial aspect of the learning process (Hattie & Timperley, 2007). Unfortunately, instructors in large-scale university courses do not have the time to provide elaborate and meaningful feedback to students. A solution could be automated feedback, a way of providing students with an elaborate feedback message tailored to their performance. Research on various forms of automated feedback shows that it can help students improve their adacemic performance. In the current study, I investigated the effects of automated email feedback on academic performance and dropout rates across three large-scale university courses. Students participated in a midterm exam after which they received only a grade (control group) or a grade with additional automatically generated feedback based on their performance. Results demonstrate that students who received this feedback did not perform better on their final exam. Additionally, results suggest that drop-out rates were not lower for students who received feedback after their midterm exam.

Keywords: automated feedback, academic achievement, dropout

The Effects of Automated Email-Feedback on Final Exam Grades and Dropout Rates

Educators and scientists agree that feedback is an important aspect of the learning process (e.g. Hattie & Timperley, 2007; Koenka et al., 2019). However, instructors in large-scale university courses do not have the time or resources to give elaborate feedback to hundreds of students throughout the semester (Glover & Brown, 2006). This unfortunately means that hundreds of students are missing out on an important aspect of their learning process. One solution to this problem has been the implementation of midterm exams, the grades of which can serve as formative assessment (Jensen & Barron, 2014; Keus et al., 2019). With the advancement of technology, it has now also become possible to add elaborate feedback to a midterm exam by using automated feedback methods.

Roediger and Karpicke (2006) have suggested that midterm exams work best for formative assessment when feedback is “elaborate and meaningful”, a description that does not apply to a letter or number grade. Grades alone may not give students enough information to improve their knowledge and skills optimally (Cain et al., 2021; Koenka et al., 2019). Additionally, receiving only a grade without any concrete advice on how to possibly improve that grade may be detrimental to students' motivation (Chamberlain et al., 2018; Koenka et al., 2019). A recent meta-analysis by Koenka and colleagues (2019) reveals that feedback through comments has a more positive effect on both achievement and motivation than feedback through grades. Similarly, a mixed-methods study by Chamberlin and colleagues (2018) found that university students who received a narrative evaluation with actionable feedback were more motivated than students who received grades.

Several studies confirm that elaborate feedback has the same positive effect in the context of midterm exams. Sato and colleagues (2018) demonstrated that feedback comments from the instructor on midterm exam answers helped students to give higher-quality answers on similar questions during the final exam. Wojcikowski and Kirk (2013) investigated multiple-choice midterms throughout a biomedical course. Students received either an elaborate exam key (that explained why certain answers were correct or incorrect) or a simple exam key (that only provided the correct answers). The researchers found that students who received the elaborate exam key performed significantly better on the final exam compared to students who received a simple exam key. In summary, elaborate feedback helps students improve their academic performance and overall motivation much more than grades do.

Though the effect of elaborate feedback is usually measured in terms of grade improvement, it may have other, less researched, positive effects. Decades ago, Tinto (1975) named feedback as an important factor in dropout prevention. This is in part because feedback is important for academic success, one of the strongest predictors of persistence (Casanova et al., 2018). Additionally, feedback can improve motivation, which also plays a major role in persistence (Kehm et al, 2019). Positive feedback is inherently motivating because it increases self-efficacy, the belief that one is competent enough to complete a goal (Burgers et al., 2015; Peifer et al., 2020). Negative (elaborate) feedback can be motivating because it informs the student about the ways in which they can reduce the discrepancy between their goal performance and their actual performance. In the context of midterms, for example, a low grade may be very demotivating, but being provided with concrete instructions on how to improve the grade may lessen the negative impact of the grade (Fong et al., 2019). Fong and colleagues (2019) found in their review of 78 studies that negative feedback that was concrete, criterion-based and delivered in person did not lead to decreases in motivation compared to no feedback. When compared to neutral feedback, criterion-based negative feedback even had a positive effect on motivation. However, Fong and colleagues (2019) also mention that negative feedback, if not concrete, criterion-based or delivered in person, may have a negative impact on motivation. Other researchers have similarly suggested that negative feedback, in some circumstances, can decrease students' motivation

(Brockner, 1987; Kim & Lee, 2019; Shin et al., 2022). If a students experience negative feedback as demotivating, it may not help to prevent dropout. As far as I know, the effects of feedback on dropout rates has not yet been investigated in prior research.

In light of the known and potential benefits of elaborate feedback, it is very unfortunate that instructors do not have the time to provide elaborate feedback in large-scale university courses (Glover & Brown, 2006). A solution could be the implementation of automated feedback, which would save instructors time while still delivering informative feedback to students tailored to their midterm performance. The term 'automated feedback' refers to a feedback message (often prewritten by the instructor) that is selected by a computer program based on a students' performance and presented to the student.

Various studies on automated feedback indicate that it can contribute positively to student learning. Cavalcanti and colleagues (2021) recently presented an overview of fourteen studies that investigated various programs to deliver feedback automatically to students. The researchers concluded from this meta-analysis that automated feedback increases student performance in learning activities. The analysis additionally found no support for manual feedback being more effective than automated feedback. Research on automated feedback specifically in the context of midterm exams seems similarly promising. Hope and Polwart (2015), for example, gave students the option to participate in pre-exam tests with automated feedback. After an incorrect answer, students would be presented with a prewritten feedback message that was tailored to the specific (incorrect) answer the student had chosen. Students who participated in more of these tests scored higher on the final exam compared to students who participated in less of these tests. However, since students' participation in these pre-exam tests depended on their own volition, this result may be confounded with student motivation. More motivated students likely put more effort into studying in general, which both makes them more likely to score high on the exam and more likely to participate in the pre-exam tests. Bulut and colleagues (2019) demonstrated that providing students with elaborate feedback on midterms through the digital score reporting program ExamVis had a positive effect on final exam grades. In this study, feedback messages were prewritten by the instructor and selected for students by the program based upon their midterm performance.

Based on these prior findings, it seems that the addition of automated feedback to midterm exams is likely to help students improve their final exam scores. However, the research about automated feedback on midterms specifically is still scarce. More research is needed in order to learn more about different automated feedback formats and about the effect of automated feedback on other outcomes besides academic achievement. These are the aims of the current study.

Firstly, the current study will investigate a different format for automated feedback. Since automated feedback is such a broad term, the feedback itself can come in many formats Cavalcanti and colleagues (2021). It is plausible that feedback in different formats may have different effects. Therefore, more research is needed on different types of automated feedback and their effectiveness. Bulut and colleagues (2019) used a program which automatically showed students feedback right after the exam. Though this saves a lot of time once it is set up, professors may not have access to this software or they may not have the time to implement new software in their course. An easier method to implement automated feedback would be to use automated emails. It is known that students perceive email feedback as high quality and as useful (Keil & Johnson, 2002). However, since such emails are text-heavy and are not viewed immediately after the exam, the effect of email feedback may be different from formats previously investigated (Bulut et al., 2019; Cavalcanti et al., 2021). The effectiveness of email feedback is more dependent on students’ motivation to open and read it than, for example, automated feedback that appears immediately after completing a test

(Bulut et al., 2019). On the other hand, emails enable students to engage with the feedback on their own time when they are in a good headspace for it, which may increase the likelihood of accepting and using the feedback (Winstone et al., 2016).

Secondly, there is still very little known about the effects of automated feedback on outcomes that are not the final grade. Dropout rates, for example, have not been studied in relation to feedback specifically as far as I am aware. Yet, feedback has been shown to be an important factor in motivation and academic performance, which in turn are known to be related to dropout (Casanova et al., 2018; Kehm et al, 2019)

The aim of the current study is to investigate the effects of automated email- feedback after a midterm exam on students' final exam performance and on dropout rates. This leads to two research questions:

1. Does automated email feedback improve students' academic performance in large-scale university courses?

2. Does automated email feedback reduce dropout rates in large-scale university courses?

A substantial amount of research has demonstrated that elaborate (automated) feedback has a positive effect on academic performance (Bulut et al., 2019; Cavalcanti et al., 2021; Chamberlin et al., 2018; Hope & Polwart, 2015; Koenka et al., 2019; Sato et al., 2018; Wojcikowski & Kirk, 2013). This leads to the hypothesis that students who received elaborate automated feedback after midterms will score better on final exams compared to students who did not receive automated email feedback after midterms.

Both positive and negative feedback can provide useful information and improve motivation (Burgers et al., 2015; Fong et al., 2019; Peifer et al., 2020). This leads to the hyothesis that students who received elaborate automated feedback after midterms will be less likely to drop out of the course compared to students who did not receive automated email feedback after midterms. However, negative feedback may decrease motivation in some circumstances (Brockner, 1987; Fong et al., 2019; Kim & Lee, 2019; Shin et al., 2022). In the case that many students perform poorly on the midterm and receive mainly negative feedback, drop-out rates may be higher among students who received automated feedback.

Research on midterm interventions consistently finds that highest and lowest-performing students do not benefit as much as average-performing students do (Dabbour 2021; Keus, 2019). For the highest-performing students this may be due to a ceiling effect: when your initial score is very high, there is not much room for improvement. Additionally, high-performing students may not make use of optional interventions because they are already meeting the standard for passing the course (Dabbour, 2021; Locke & Latham, 1990). Low-performing students may also be unlikely to use feedback, though they theoretically have the most to gain from it. Students may attribute their low midterm score to simply not trying and may not read the feedback, expecting it will not give them additional information. Additionally, a lot of negative feedback may discourage students (Brockner, 1987; Fong et al., 2019; Kim & Lee, 2019; Shin et al., 2022), leading them to reject the feedback message. Therefore, students' midterm grades will be taken into account when investigating the relationship between automated email-feedback and academic performance or dropout.

Method

Design

This study was an intervention study using a randomized controlled trial design. Students were divided based on gender and grouped based on their high school final exam grade in mathematics, to make sure no large differences in mathematical ability would be present between conditions. Within these groups, students were randomly assigned to one of eight conditions. Depending on the condition, students would receive either a grade and feedback on their midterm or only a grade. The sorting into conditions was repeated for each course. For every student, whether they were part of the experimental condition (grade +

feedback) or the control condition (grade only) could therefore be different for each course. This allows for comparison of effects between courses and helps rule out possible learning effects.

Participants

Participants were first-year students from the Bachelor track Business Administration at the University of Groningen who started the program in September 2016. Students were first notified about the study during their introduction to the program on September 5 of 2016, one week before the start of course lectures. In this introduction, students were not informed about the specific aim of the study. They were told that the aim of the study was to develop ways to support students' studying and lecturers' teaching. On September 8, 2016, students received an invitation to participate in the study through their mentor. The invitation contained general information about the topic and procedure of the study and a request for students' informed consent to participate in the study. Additionally, students were informed that giving permission would mean that researchers will be able to link their student record to the outcomes of the experiment. Mentors like the students, were only aware of the general goal of the study.

Ethical Considerations

Students' data was collected from three courses from the Bachelor track Business Administration at the University of Groningen: Management Science, Statistics and Supply Chain Operations. The same cohort of students participated in all three courses. Out of 318 students, 280 students gave informed consent for their data to be used for study.

This study was approved by the ethics committee of the faculty of Economics & Business (FEB-20190410-7909).

Procedure

For each course, midterm exams were held at the halfway point of the course. After each midterm exam students filled out a short survey, the data of which was not analyzed in the current study.

Midterms and exams in the Management Science and Supply Chain Operations courses consisted of a combination of multiple choice and essay type questions. Midterms and exams in the Statistics course consisted of essay type questions. The course instructors created midterm exams by designing questions based on learning objectives for the course. Each midterm exam assessed six to ten learning objectives. Each learning objective was represented by an average number of 4 questions. Using the scores that students achieved on these questions, researchers then determined whether students scored high, moderate or low on a certain learning objective.

Feedback messages were prewritten by the researchers for low, average and high scores on each learning objective on the midterm exam (See Appendix A for an example). For each learning objective, a separate message was composed for each score level. A student who scored high on a learning objective would not receive additional feedback apart from the statement that they had achieved a high score on this learning objective. A student who scored moderate would receive some additional suggestions for study and practice and a student who scored very low would receive even more additional suggestions for study and practice of the topic. Feedback messages were designed based on the properties of effective feedback described by Hattie & Timperley (2007). Excel was used to automatically select the appropriate feedback messages based on a student's score on each learning objective and, after that, combine the feedback messages on each learning objective into an email.

Grades and feedback were delivered in a standardized email to students' university email approximately one week after the completion of the midterm exam. The control group received only a grade via email, at the same time.

At the end of the semester, there was a final exam for each course. After the final exam students again filled out a short survey, the data of which was not analyzed in the current study.

Variables and Instruments

Academic performance was measured using students' final exam grades for each course. Midterm grades and final exam grades were independent, that is, midterm grades did not contribute to final exam grades. Midterm grades did contribute to the final grade of the course. Both midterms and final exams in all courses were scored on a scale of 0 to 10 with one decimal (for example, 5,4 but not 5,43). The threshold for passing each exam was set at a grade of 5,5.

Dropout was operationalized as final exam attendance. If a student did not attend the final exam, they were assumed to have dropped out of the course.

Data Analysis

Data was analyzed using IBM SPSS Statistics (Version 28). To assess the direction and strength of the relationship between automated email feedback and academic performance, linear regression was used. To predict the likelihood of dropout from feedback condition (automated feedback or no feedback), logistic regression was used.

As discussed prior, whether students make use of midterm feedback may be dependent on their performance on the midterm. Perhaps an effect of feedback in the general group could be obscured by an interaction between midterm performance and feedback use. Because of these considerations, midterm grades and their interaction with feedback condition were included in both regression analyses.

Results

This section begins with an exploration of the data to provide relevant context before presenting the results of the regression analyses. First, participant demographics can be found in Table 1. Next, assumption checks are presented before continueing to an exploration of the midterm exam results. After this exploration of the data, the regression analyses are presented.

Table 1

Participant Demographics

Course	Group	N	Male	Female	Mean age in years (SD)
	Feedback	143	102	41	18.76 (1.222)
Management Science	Control	146	101	45	18.70 (1.478)

	Total	289	203	86	18.73 (1.335)
	Feedback	128	87	41	18.52 (.988)
Statistics	Control	142	97	45	18.68 (1.286)
	Total	270	184	86	18.60 (1.155)
	Feedback	128	84	44	18.44 (.911)
Supply Chain Operations	Control	158	110	48	18.63 (1.274)
	Total	286	195	91	18.55 (1.128)

Assumption Checks

Data from each course showed no significant deviations from normal distribution or other significant assumption violations. Some outliers were present in the final exam grades, but these did not significantly affect the analysis results.

Midterm Performance

In all three courses, all students attended the midterm exams. Midterm exams were scored on a scale of 1 to 10 with one decimal (e.g. 5.5, but not 5.54). The threshold for passing the exam was set at a grade of 5.50. Across all courses, average midterm grades and grade distributions were similar for students who would receive feedback after the midterms and students who would not receive feedback (Table 2). This suggests that students in both conditions were comparable in terms of academic achievement.

Table 2

Mean Midterm Exam Grades for Each Condition per Course

Course	Group	M (SD)
Management Science	Feedback	6.02 (1.718)
	Control	6.15 (1.584)
Statistics	Feedback	5.44 (1.395)

	Control	5.63 (1.489)
Supply Chain Operations	Feedback	5.54 (1.623)
	Control	5.40 (1.655).

Final Exam Grades

Students who did not attend the final exam were not included in the following analysis of final exam grades, since they did not receive a final exam grade. See Table 5 for attrition numbers for each condition, per course. The remaining Management Science sample consisted of 191 male and 84 female students with a mean age of 18.69 years (SD = 1.338). The remaining Statistics sample consisted of 172 male and 83 female students with a mean age of 18.60 years (SD = 1.173). The remaining Supply Chain Operations sample consisted of 179 male and 86 female students with a mean age of 18.51 years (SD =.946).

Grades on the final exam were scored on a scale of 1 to 10 with one decimal. The threshold for passing the exam was set at a grade of 5.50.

Descriptives

For Management Science, final exam scores were lower on average than midterm exam scores. In the Statistics course and in the Supply Chain Operations course, students seem to have done better overall on the final exam compared to the midterm exam (Table 3). Across all courses, students in the feedback condition seemed to score lower overall on the final exam compared to students in the control condition, but the difference is very small.

Table 3

Mean Final Exam Grades for Each Condition per Course

Course	Group	M (SD)
Management Science	Feedback	5.78 (1.366)
	Control	5.83 (1.264)
Statistics	Feedback	5.97 (1.984)
	Control	6.02 (1.945)

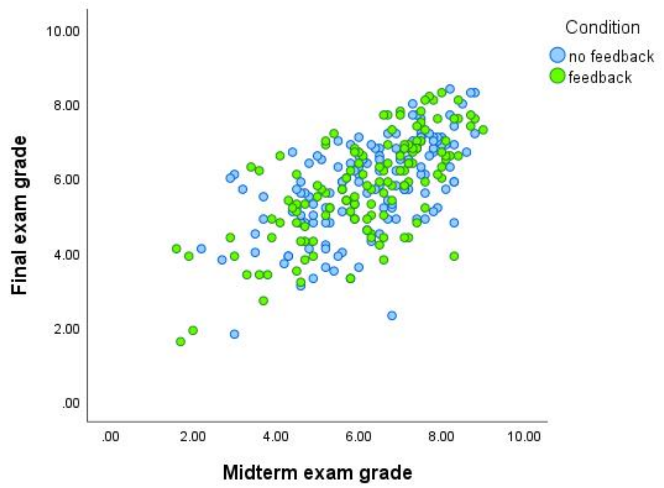
Supply Chain Operations	Feedback	6.16 (1.560)
	Control	6.24 (1.400)

Although the mean grades suggest a small, but consistent difference between conditions, the scatterplots in Figure 1 show that the distribution of final exam grades is very similar between the feedback condition and the control condition.

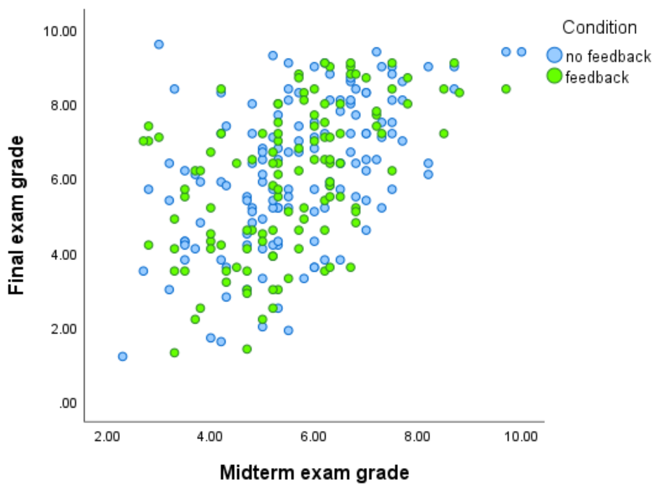
Figure 1

Scatterplots of Final Exam Grades for Each Course

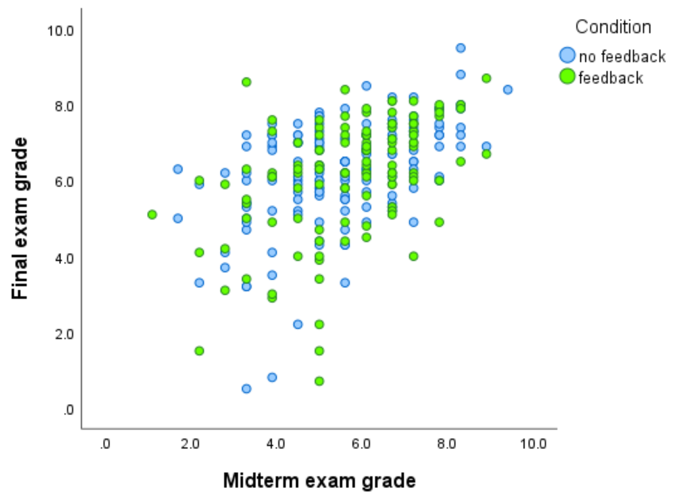
A. Management Science



B. Statistics



C. Supply Chain Operations



Regression Results

For each course, a multiple regression was run to predict final exam grades from midterm exam grades, feedback condition and the interaction between midterm exam grades and feedback condition. All models were significant ($p < .001$, see Table 4 for F values and degrees of freedom). However, only midterm exam grades significantly predicted final exam grades.

Table 4

Overview of Important Regression Results per Course for Final Exam Grades

Course	*F*	df (between, within)	Nagelkerke R^2	Significant predictors	*p*
Management Science	63.497	3, 271	.41	Midterm grade	<.001
Statistics	28.692,	3, 251	.26	Midterm grade	< .001.
Supply Chain Operations	26.812	3, 261	.24	Midterm grade	< .001.

Note. p refers to the p-value of the significant predictor, which is in this case midterm grade.

Dropout

Dropout was operationalized as 'not attending the final exam for the course'. See a summary of dropout numbers for each condition per course in Table 5.

Table 5

Attrition (Dropout) Statistics for Each Condition, Per Course

Course	Group	N	Dropped out	Attended final exam
	Feedback	143	9	134
Management Science	Control	146	5	141
	Total	289	14	275
	Feedback	128	10	118
Statistics	Control	142	5	137

	Total	270	15	255
	Feedback	128	10	118
Supply Chain Operations	Control	158	11	147
	Total	286	21	265

Note. N represents the number of students who consented to participate in the study and attended the midterm exam.

Assumption Checks

Data from each course showed no significant assumption violations.

Regression Results

For each course, a logistic regression was performed to analyze the effects of feedback condition, midterm grade and the interaction between the two on the likelihood that students would drop out before the final exam. All models were statistically significant (p < .001, see Table for Chi-squared and p-values). However, midterm grade was the only significant predictor of dropout likelihood in all three courses (Table 6).

Table 6

Overview of Important Regression Results per Course for Dropout

Course	$\chi 2$ (df)	Nagelkerke R^2	Significant predictors	p
Management Science	33.480 (3)	0.34	Midterm grade	.003
Statistics	19.591 (3)	0.30	Midterm grade	.007
Supply Chain Operations	21.719 (3)	0.18	Midterm grade	.007

Note. p refers to the p-value of the significant predictor, which is in this case midterm grade.

Discussion

In this study I examined whether automated email feedback after midterm exams could 1) improve students' academic achievement and 2) reduce dropout rates in three large-scale university courses. First, I summarize the results and discuss them in relation to previous research, after which possible reasons for these findings are explored. This is followed by a discussion of limitations and suggestions for future research.

Automated Email Feedback and Grades

Automated email feedback after a midterm exam did not improve students' final exam grades across three large-scale university courses. Results showed that only midterm grades predicted final exam grades, so that students with higher midterm grades also scored higher on the final exam and students with lower midterm grades also scored lower on the final exam. This is consistent with previous findings on midterm interventions by Jensen and Barron (2014), who found that midterm exam grades strongly predict final exam grades. It is surprising, however, that even with informative automated feedback students were not able to improve their academic performance. This is in contrast with prior findings on automated feedback (Cavalcanti et al., 2021; Hope & Polwart, 2015; Bulut et al., 2019), especially Bulut and colleagues (2019) who found that students did improve their final exam grades after participating in midterm exams with automated feedback.

Automated Email Feedback and Dropout

Automated email feedback did not reduce dropout rates across three large-scale university courses. Only midterm grades were found to be related to students' likelihood of dropping out, with lower midterm grades predicting a higher likelihood of dropping out. This is in line with prior research on dropout, which has found that academic achievement (often in terms of grades) is one of the strongest predictors of dropout (Casanova et al., 2018). It seems that, in the current study, the elaborate feedback accompanying low midterm grades did not work to mitigate the effect of a low grade on students' intention to drop out.

Exploration of the Lack of Feedback Effect

It is possible that providing a grade along with the feedback may have undermined the effectiveness of the feedback. Students may have felt that feedback did not add much information to what they could deduct from the grade. Students who scored very low on the midterm may have felt they already knew that they had to improve 'everything'. Students who passed or scored high similarly may not have read the feedback as they already reached the standard. Perhaps, then, the feedback in this study would have been more effective if it had not been accompanied by a grade. Studies that found positive effects of elaborate feedback have often provided students with elaborate feedback only, without a grade (Chamberlin et al., 2018; Koenka et al., 2019). Indeed, Keuper-Tetzel and Gardner (2021) confirmed that delivering elaborate feedback before giving a grade encourages students to read and implement the feedback.

However, it cannot be concluded with certainty that students did not use the feedback.

Students' Reception and Implementation of Feedback

The current study included no measure of students' use of the feedback. It is assumed that students read the feedback email because it contained their grade, but it is uncertain if students 1) fully read through the feedback and 2) implemented the feedback effectively into their studying practice for the final exam.

Prior feedback research has also encountered this problem. Even when students ask for feedback and indicate that they are willing to use it, many students do not end up using it (Bulut et al., 2019; Daniels & Bulut, 2019; Winstone et al., 2017). Bulut and colleagues (2019), for example, report in their first experiment that only 30% of students viewed both of the automated feedback score reports offered to them. Out of all students, 35% did not view any of the automated feedback offered to them. This is in contrast to 97% of students indicating that they would be willing to use feedback on their midterm exam to study for their final exam (Bulut et al., 2019). Daniels & Bulut (2019) similarly report that students judged elaborate feedback reports as useful, but this was not related to improved academic performance.

Unfortunately, the current study demonstrates anew that low-performing students, who arguably need feedback the most, are unlikely to benefit from feedback (Jensen & Barron, 2014). Low-performing students may struggle with underlying problems in motivation or self-regulation and this may cause them not to use the available feedback (Locke & Latham, 1990). When given the chance to improve their performance with an optional second midterm combined with grade dropping, more of the lowest-performing students would not participate in the midterm compared to the average or slightly low performing students (Dabbour, 2021). This may have been due to a general lack of motivation and self-regulation skills among the lowest-performing students which led them to score very low on the first midterm exam and concurrently led them to decide not to rescue their grade with a second exam.

Considerations on the Complexity of Dropout

Though the current sample was large, there was a relatively small amount of dropout in both conditions. This may have made it difficult to detect an effect of feedback on dropout rates in the first place.

This possibility aside, the lack of an effect of automated email feedback on dropout rates is not necessarily surprising. Although it has been suggested that elaborate feedback is an important variable in student persistence (Tinto, 1975), dropout is a deceptively complex issue with many contributing factors (Casanova et al., 2018, Kehm et al., 2019). In the current study I was not able to keep track of the other factors and their complex relationships in affecting the likelihood of dropout. Even if a positive effect of feedback on dropout may have existed in the study, it may have been obscured or nullified by effects from other factors such as social and academic integration, motivation and sociodemographic background of students (Kehm et al., 2019).

Additionally, not all dropout may be preventable. A few students who dropped out had scored quite high on the midterm exam. It is unlikely that in their case dropout was due to a lack of self-efficacy or not knowing how to improve, which could be mitigated by the feedback. Rather it seems plausible that other factors were at play here such as motivation or perhaps a change in major. Finally, perhaps we should not aim to prevent all dropout. In fact, Jensen and Barron (2014) suggest that dropping out may be the right choice for students when they really struggle to meet the demands of certain courses. It may be more beneficial for them to realize early that a certain academic path is not suited for them so they can choose something else without losing much time or money to a degree they may continue to struggle in.

Limitations

The main limitation of the current study is that it can not provide insight in students' implementation of the automated email feedback. From the current data it is impossible to deduce whether the feedback intervention truly reached (all) students. This makes it difficult to draw strong conclusions from the lack of feedback effect.

Additionally, the relatively small number of students who dropped out in each condition may not have been enough to show an effect of feedback.

Furthermore, the time to process and implement the feedback may not have been enough for all students to show significant improvement in their academic performance. Students received feedback approximately one week after the midterm exams, which took place in the middle of the semester. This means students only received the feedback when they were already more than halfway through the semester. By this time, it may have been difficult for students, especially low-scoring students, to implement all the feedback before the exam.

Finally, I was not able to investigate whether automated email feedback may have had different effects for different exam subtopics. Concrete, actionable feedback was provided for all learning objectives, but perhaps feedback was more useful or easier to implement for some learning objectives than others. Learning objectives that require abstract thinking, for example, may require in-person explanation based on the student's level of understanding. Such learning objectives may benefit less from advice to reread corresponding slides compared to objectives that require mainly memorization of theory.

Suggestions for Future Research

In order to investigate more thoroughly whether email is an ineffective format for automated feedback, a measure of student implementation of feedback must be included in future research on automated email feedback. It seems there exist barriers to students' implementation of feedback. Researchers are already identifying a number of these barriers

(Winstone et al., 2017), but future research can hone in on barriers that arise from automated feedback specifically. Once more is known about students' reception and implementation of email feedback, automated email feedback on midterms may be compared to other automated feedback formats that have been proven effective (Bulut et al., 2019).

Furthermore, researchers have suggested that, instead of aiming to deliver "more" feedback, it may be helpful to focus on ways to increase students' engagement with the feedback (Hepplestone et al., 2011; Winstone et al., 2017). Automated feedback and online learning programs can be of assistance, for example by setting a requirement to respond to feedback (Hepplestone et al., 2011). It would be interesting to investigate similar 'engagement-encouraging' features with automated email feedback.

In order to detangle the effects of feedback on dropout, I believe that qualitative methods will be useful. Dropout is a very complex issue, and hearing from dropped out students themselves may shed more light on the factors that play a role. Future research on the link between feedback and dropout should take into account as many factors as possible that may affect the link between feedback and dropout risk.

Conclusion

Automated email feedback after a midterm exam does not seem to help students improve their academic performance in general. However, this form of feedback also does not cost much effort to set up. Instructors may decide that the potential of automated email feedback helping a handful of motivated students is worth this small extra effort. As for now, I believe that too much is still unknown about students' interaction with automated email feedback to make a strong case for or against its implementation in large-scale university courses.

References

Bulut, O., Cutumisu, M., Aquilina, A. M., & Singh, D. (2019). Effects of digital score reporting and feedback on students' learning in higher education. *Frontiers in Education, 4*. https://doi.org/10.3389/feduc.2019.00065

Borrella, I., Caballero-Caballero, S., & Ponce-Cueto, E. (2019). Predict and intervene. Proceedings of the Sixth (2019) ACM Conference on Learning @ Scale. https://doi.org/10.1145/3330430.3333634

Brockner, J., Derr, W. R., & Laing, W. N. (1987). Self-esteem and reactions to negative feedback: Toward Greater Generalizability. *Journal of Research in Personality, 21*(3), 318–333. https://doi.org/10.1016/0092-6566(87)90014-6

Burgers, C., Eden, A., van Engelenburg, M. D., & Buningh, S. (2015). How feedback boosts motivation and play in a brain-training game. *Computers in Human Behavior, 48*, 94–103. https://doi.org/10.1016/j.chb.2015.01.038

Cain, J., Medina, M., Romanelli, F., & Persky, A. (2021). Deficiencies of traditional grading systems and recommendations for the future. *American Journal of Pharmaceutical Education, 86*(7), 8850. https://doi.org/10.5688/ajpe8850

Chamberlin, K., Yasué, M., & Chiang, I.-C. A. (2018). The impact of grades on student motivation. *Active Learning in Higher Education*, 24(2). https://doi.org/10.1177/1469787418819728

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Casanova, J. R., Cervero, A., Núñez, J. C., Almeida, L. S., Bernard, A. (2018). Factors that determine the persistence and dropout of university students. *Psicothema, 30* (4), 408-414 https://doi.org/10.7334/psicothema2018.155

Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y.-S., Gašević, D., & Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, *2*(100027). https://doi.org/10.1016/j.caeai.2021.100027

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and Quantitative Synthesis. *Psychological Bulletin, 132*(3), 354–380. https://doi.org/10.1037/0033-2909.132.3.354

Dabbour, E. (2021). Motivating engineering students by providing two midterm exams and dropping the lower mark. *Journal of Civil Engineering Education, 147*(1). https://doi.org/10.1061/(asce)ei.2643-9115.0000024

Fong, C. J., Patall, E. A., Vasquez, A. C., & Stautberg, S. (2018). A meta-analysis of negative feedback on intrinsic motivation. *Educational Psychology Review, 31*(1), 121–162. https://doi.org/10.1007/s10648-018-9446-6

Gifford, T. (2023). Improving Nursing Student Retention: Early Identification of Determinants Causing Attrition in Nursing Academic Programs. *Doctor of Nursing Practice Final Manuscripts, 246*. https://doi.org/10.22371/07.2023.039

Glover, C., & Brown, E. (2006). Written feedback for students: Too much, too detailed or too incomprehensible to be effective? *Bioscience Education, 7*(1), 1–16. https://doi.org/10.3108/beej.2006.07000004

Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77*(1). https://doi.org/10.3102/003465430298487

Hope, S. A., & Polwart, A. (2012). Engagement with online pre-exam formative tests improves exam performance and feedback satisfaction. *Bioscience Education, 20*(1), 37–52. https://doi.org/10.11120/beej.2012.20000037

IBM Corp. Released 2021. IBM SPSS Statistics for Windows, Version 28.0. Armonk, NY: IBM Corp

Jensen, P., & Barron, J. (2014). Research and teaching: Midterm and first-exam grades predict final grades in biology courses. *Journal of College Science Teaching, 044*(02). https://doi.org/10.2505/4/jcst14_044_02_82

Kehm, B. M., Larsen, M. R., & Sommersel, H. B. (2019). Student dropout from universities in Europe: A review of empirical literature. *Hungarian Educational Research Journal, 9*(2), 147–164. https://doi.org/10.1556/063.9.2019.1.18

Keus, K., Grunwald, J., & Haave, N. (2019). A Method to the Midterms: The Impact of a Second Midterm on Students' Learning Outcomes. *Bioscene: Journal of College Biology Teaching, 45*(1), 3-8.

Kim, E. J., & Lee, K. R. (2019). Effects of an Examiner's positive and negative feedback on self-assessment of skill performance, emotional response, and self-efficacy in Korea: A quasi-experimental study. *BMC Medical Education, 19*(1). https://doi.org/10.1186/s12909-019-1595-x

Koenka, A. C., Linnenbrink-Garcia, L., Moshontz, H., Atkinson, K. M., Sanchez, C. E., & Cooper, H. (2019). A meta-analysis on the impact of grades and comments on academic motivation and achievement: A case for written feedback. *Educational Psychology, 41*(7), 922–947. https://doi.org/10.1080/01443410.2019.1659939

Kuepper-Tetzel, C. E., & Gardner, P. L. (2021). Effects of temporary mark withholding on academic performance. *Psychology Learning & Teaching, 20*(3), 405–419. https://doi.org/10.1177/1475725721999958

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Prentice Hall.

McIntyre, S. H., & Munson, J. M. (2008). Exploring cramming. *Journal of Marketing Education, 30*(3), 226–243. https://doi.org/10.1177/0273475308321819

Peifer, C., Schönfeld, P., Wolters, G., Aust, F., & Margraf, J. (2020). Well done! effects of positive feedback on perceived self-efficacy, flow and performance in a mental arithmetic task. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.01008

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning. *Psychological Science, 17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Sato, B. K., Dinh-Dang, D., Cruz-Hinojoza, E., Denaro, K., Hill, C. F., & Williams, A. (2018). The impact of instructor exam feedback on student understanding in a largeenrollment biology course. *BioScience, 68*(8), 601–611. https://doi.org/10.1093/biosci/biy058

Shin, D. D., Kim, S., Lee, M.-J., Jiang, Y., & Bong, M. (2022). Role of perceived competence and task interest in learning from negative feedback. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.830462

Spitzer, H. F. (1939). Studies in retention. *Journal of Educational Psychology, 30*(9), 641–656. https://doi.org/10.1037/h0063404

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89–125. https://doi.org/10.3102/00346543045001089

Winstone N., Nash, R., Rowntree, J. & Parker, M. (2017) It'd be useful, but I wouldn't use it: barriers to university students' feedback seeking and recipience, *Studies in Higher Education, 42*(11), 2026-2041, https://doi.org/10.1080/03075079.2015.1130032

Wojcikowski, K., & Kirk, L. (2013). Immediate detailed feedback to test-enhanced learning: An effective online educational tool. *Medical Teacher, 35*(11), 915–919. https://doi.org/10.3109/0142159x.2013.826793

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin, 147*(4), 399–435. https://doi.org/10.1037/bul0000309

Appendix A

Example of Automated Email Feedback

From: "Management Science FEB" <managementscience@rug.nl>
To: X2@student.rug.nl
Date: 10/12/2016 3:06:00 PM
Subject: Results midterm Management Science

Dear X2,

The results of the midterm of Management Science (EBP025A05), date Monday October 10, are known.

Your result on the midterm: 3,9

In this email you find personalized feedback.

You have scored high on the next topics:

• Model a problem in mathematical notation with respect to a forecasting model.
• Interpret model results in the context of the original problem.

You have received a moderate score on the following topics, hence we see room for improvement:

• Active usage of mathematical notation with respect to summations. We advise to develop routine by making somewhat more challenging assignments, for example the practice material of week 37 and the homework assignment of week 38.
• Model a problem in mathematical notation with respect to the definition of variables and parameters. Slide 17 of the tutorial in week 39 gives a correct explanation of the use of parameters in assignments. The lecture in week 39 (slide 11 and 12) give further explanation on the definition of decision variables.
• Excel skills (i.e., formulas). Develop routine by practising with the assignments in building Excel formulas: tutorial week 37 question 2 and 4, tutorial week 39 question 3j-3l. Give specific attention to preventing carelessness errors.

You have scored very low on the following topics, hence we conclude that improvement is necessary:

• Active usage of mathematical notation with respect to matrices. The basics of matrix mathematics has been explained in a video tutorial and on the slides of tutorial week 39 (slides 28-54). The additional theory, mathematical notation Section 3, is important to understand all rules with respect to matrix mathematics. It is important to develop sufficient routine, the tutorial exercises and practice material of weeks 39 and 40 can support you in this process.
• Model a problem in mathematical notation with respect to the definition of functions and constraints. The first assignment in tutorial week 39 shows how functions and constraints are constructed, hence reconsider these instructions. Try to build routine with the practice material of week 39.
• Interpret model results using Excel output. Study specifically slide 20 of the lecture in week 37. And practice the practical of week 37 as this includes regression analysis.

The answer model is published on Nestor under course documents > Midterm. The grade center will show your results under ‘My Grades’.

Best regards,
Jon Hummel
Coordinator Management Science