

ATOMIZED FEEDBACK, AN APPROACH TO AUTOMATIC FORMATIVE ASSESSMENT

Arild Steen, Hans Richard Movik

UiT – The Arctic University of Norway

Abstract

In this article, we describe an approach to give automatic formative evaluation by atomizing the feedback. The mechanics in this approach is a large number of multiple-choice tests. The multiple-choice tests are used as voluntary assignments and gives controlled progression by releasing the tests in an order. Instruction is in the spirit of flipped classroom with short video lectures accompanied by one to several above-mentioned assignments. The preliminary results show reduced dropout, higher student satisfaction, and less fail on final exam.

The background for this approach submerged from a multi-campus/online course with 300+ students. The aim was to create the same environment for learning for all categories of students, on campus, remote campus and online students.

Atomized formative evaluation is achieved by giving feedback down to each individual multiple-choice answer. The multiple-choice questions are in the nature of practical tasks where students are expected to use external tools, calculations and simulation to find answers. The multiple-choice test collects the students' findings. The number of correct answer gives a summative evaluation. After turning in the test, answers are available together with feedback for each answer. Correct answers explain why this is correct and a wrong answer explains why this is wrong in addition to how and why the student should go by finding the correct answer. The questions are designed with no "dummy" answer, so when choosing a wrong answer, it is a plausible answer, and an explanation for this is given as feedback.

The goal is to encourage the student to try again both to improve the summative score and to fill the gap in knowledge. When forming the feedback, one must consider both the type of question and the choices made available.

Using atomized feedback improves the formative evaluation greatly by being instant, continuous, to the point and informative. This is evaluation for learning that both guides and encourage the student. In addition, since instructors/TA do not use time to evaluate results on assignments, they can guide / tutor individual students who need follow up.

Student performance is evaluated by achieved grade, student satisfaction, progression, and the number of assignments compared to grade. The preliminary findings compared with earlier student performance in the same course but without atomized formative feedback show a reduced number of dropouts and reduced number of students failing the final exam.

Keywords: Automatic formative evaluation, flipped classroom, multiple campus / online students

1 BACKGROUND

The Arctic University of Norway has campus in Tromsø, Narvik, Harstad, Bodø, Mo, Alta, and Kirkenes. Faculty for Engineering Science and Technology at Campus Narvik is the hub for teaching and research in different engineering disciplines. Since the university has a multi campus organization, teaching in several courses requires online participation at the other campuses. One such course is a general engineering course, Programmable Calculation and Statistics. This course has an average of 300+ students located at several campuses and about 1/3 of the students are individual online students.

This course has two parts: Programmable Calculations, and Statistics. Each part is on five credits. There are considerable cooperation and collaboration between the two regarding subject matters. However, the approach described in this article relates for the most to the teaching in the Programmable Calculations part of this course.

The course has limited resources that consist of:

- 1,5 teacher resource
- 5-6 student assistant

- Computer lab with 50 seats (at best)
- The largest auditorium with 300 seats (only used at startup)

The largest student group is at campus Narvik with about 180 students, Campus Bodø has about 20, then there are 20 on campus Alta, and a few at other campus and finally about 90 online students from all over Norway.

Since the course is thought on several campuses, has about 1/3 online students (both synchronous and asynchronous), and has limited resources, traditional lecture methods do not suffice. We had to come up with new methods and this approach is described in the following chapters.

2 FORMATIVE ASSESSMENT

Our approach has its roots in formative assessment (aka, “assessment for learning”). There are several publications regarding the method. Randy Elliot Bennett goes into the issues regarding effectiveness, domain consideration, measurement principles, teacher-support entailment and concludes among other that “conceptualizations (of formative assessment: ed.) should incorporate fundamental measurement principles that encourage teachers and students to recognize the inferential nature of assessment.” Moreover, that “formative approaches should be conceptualized as part of a comprehensive system in which all components work together to facilitate learning” [1]. We have recognized that there are different views on the method and we have looked further into the teacher-support issue by approaching this in an attempt to liberate the teacher-support by automate the formative assessment. David J. Nicol & Debra Macfarlane-Dick reinterpreted formative assessment “to show how these processes can help students take control of their own learning, i.e. become self-regulated learners”. Since our approach have automatic feedback, students are given some control over their own learning by pacing the self through a progression controlled row of modules. One has to complete one module before given access to the next. Nicol and Macfarlane-Dick go on to “identify seven principles of good feedback practice that support self-regulation” [2]

Good feedback practice [2]:

1. helps clarify what good performance is (goals, criteria, expected standards);
2. facilitates the development of self-assessment (reflection) in learning;
3. delivers high quality information to students about their learning;
4. encourages teacher and peer dialogue around learning;
5. encourages positive motivational beliefs and self-esteem;
6. provides opportunities to close the gap between current and desired performance;
7. provides information to teachers that can be used to help shape teaching.

Our approach covers several of these practices with varying strength. #1 is strongly covered by summative and formative feedback, #2 is unsure. #3 can argue that students get some information about their learning. #4 not much dialog with teacher, uncertain how much between peers. #5 students get positive motivation and since there are many small steps (i.e. assignments) the self-esteem rises. #6 is strongly present and #7 likewise.

Bernt Arne Bertheussen presented in 2014 a concept where students solved financial problems in spreadsheets with automatic formative feedback. This was an algorithm that looked at both part-solutions and final result to aid students in the problem solving. [3] He had focus on enhanced learning outcome and increased motivation.

This kind of motivation is inherited in that we give feedback on every answer, but in addition, we have tried to give cheering by using motivating comments in grading (the summative part).

Knut Bjørkli og Ketil Arnesen used in 2013 a concept with several under way summative multiple-choice test where the individual results were used as starting point for the individual formative feedback. A control group did not participate in the under way summative multiple-choice tests concept. The final written exam showed significant better results for the group participating in the under way summative multiple-choice tests than the control group. [4] This approach lines up with the good feedback practice #7. However, in our setting we could not use this due to limited resources and it is not feasible for 1,5 teacher to have dialog with 300+ students in a semester.

Nancy Frey and Douglas Fisher writes about “...what formative assessment is all about— taking action based on student performance” [5]. In our approach, we monitor student performance based on number of completed assignments and what summative score students get. For those students where the total

“so far” score is above 75% we do nothing. For those below we contact by e-mail and ask how they do and if they need some help. For those that reply on this we give individual tutoring to sort out where the problem lies.

John Hattie and Helen Timperley notes that feedback works by reducing the discrepancy between current and desired understanding. Moreover they also notes that not all methods that are effective in enhancing learning, but students are more likely to increase effort when the intended goal is clear [6]. In our approach, each voluntary assignment is a goal, in that students should get 100% on each. They get feedback that explain what they did wrong, and how to find aid to correct. We have tried to use motivating language to cheer on to try again.

3 OUR APPROACH

The teaching mechanics of this course uses a variant of flipped classroom. With the preconditions mentioned in the Background section, traditional lectures in auditorium and exercises in computer lab would not do. We needed an approach that could work well both on and off campus. We were initially interested in using flipped classroom, but the “in-class” activities were not manageable. Online students had limited possibilities of attending and we could not even fit all the students at campus Narvik on the schedule for the computer lab.

3.1 Flipped variant

We chose to use most of what flipped classroom concept had to offer with a few changes. The students had to be able to monitor their own progress. Therefore, we were more focused on the assignments than on the lectures. We designed assignments and video lectures as a replacement for ordinary lectures. Since we have limited laboratory resources only students struggling with the assignments was invited to attend lab (but no one was refused). The voluntary assignments are either executed by the student alone at his/her own pace and time, or in the lab/class together with faculty and student assistants available. Simultaneously with the lab/class activity, the same activity is conducted online through Adobe Connect Online Meeting (ACOM) with audio and video to serve other campus and online students. It is not mandatory to participate, and ACOM with student assistants were also scheduled in the afternoon, to meet the needs from online students not able to attend the daytime labs. Access to instruction and assignments are controlled by preconditions in the module functionality in Canvas. This is sett up so that one must complete assignments with a result above a set score to get access to the nex module

All instruction and teaching is video with a length of 6 to 12 minutes. There are about 60+ instructional videos. There are about 40 voluntary assignments accompanying instructional videos. All the assignments are using multiple-choice quiz / test that collect the answers. The test question requires students to use tools outside the test i.e. MatLab, Excel, to calculate / find answer.

Feedback from students both in surveys and in direct contact with some of them were in general positive. Online students responded that this course was better suited for online study than many of the other courses offered. Both campus and online students felt that this course gave them some degree of control over their own progress and favored this. Both of this were related to two things: That all material was available online and the possibility to check their own progress.

3.2 Motivating grading

In the first edition of the course, the only feedback to the students was a percentage score on each test. The scale was divided in to three parts with a characteristic to each: <40% “Poor achievement, not approved”; <60% “Not adequate, but let go ...”; > 60 % “Approved”. The overall fail limit for the course is 40%, but on each test, it is set to 60%. This feedback was only summative. It had no formative qualities and did not encourage students to try to improve.

The first improvement was to refine the scale and give the characteristics a more encouraging touch.

- > 92% “Excellent, star material!”
- > 80% “Great achievement, approved and then some!”
- > 60% “Approved, but can be improved.”
- > 40% “Not quite adequate, try again.”
- < 40% “Achievement below goals, study some more and try again”

All though this might look better, it does not help the student to pinpoint any gaps in knowledge/skills or how to fill it. To be able to do this without coaching, each individual student we started experimenting with the possibilities in the LMS to provide feedback in questions/possible answers. We soon realized that the focus had to be on possible answers, i.e. formative assessment.

3.3 Feedback on all answers

Traditionally, students do not get much feedback on their answers in a multiple-choice quiz, only points for correct and semi-correct, and zero points for wrong answer. Students get the results after completing a test, sometimes immediate after test, other times it takes up to days to get the result due to necessary instructor intervention and evaluation (depending on question types).

In our approach, after completing the test, the system itself indicates whether a chosen answer is correct or not. Students get feedback information telling them why the answer is correct or not and they get a summative score. The feedback must have information of value to the student. This can for instance be:

For incorrect alternatives

- Explaining what and why it is wrong.
- Explaining what the student might have done wrong to obtain this answer.
- Refer to lecture, literature, example.
- Refer to help section.
- Make student aware of pitfalls.
- ...

For correct alternatives

- Explaining why it is correct. The student is assured that the concept is understood.
- Refer to common use of this approach.
- Giving extra information on the technique
- ...

The feedback on each possible answer can consist of one or more of these depending of the question itself, alternatives and where in the progress of the course this test belongs. The important thing to remember is that the student only gets feedback on the chosen alternative. Good feedback on correct alternatives are just as important as on incorrect alternatives. The overall goal of feedback is to encourage students to fill the gap in knowledge and thereby improve the summative score

Example 1: (translated from Norwegian)

Use the recipe from the example and solve the following set of equations I MATLAB:

$$2x + y = 5$$

$$x - 2y = -5$$

What is the x and y values of the solution?

Answer alternatives:

x = 1 and y = 1 (wrong answer)

Check that you have the correct numbers in the matrices, and that they are in the right place.

x = 1 and y = 3 (the correct answer)

You have used the recipe correctly. Using $A \setminus B$ or $\text{inv}(A) * B$ will both give the correct answer.

x = 3 and y = 1 (wrong answer)

Check which element of the solution is the x value and which is the y value.

You get an error message. There is no unique solution to the problem (wrong answer)

Be sure you have followed the recipe from the example thoroughly. Check that you have used the correct operators/signs.

Example 2: (translated from Norwegian)

Starting point in Excel sheet named "Friction". You have your obstinate little brother that "do not want to..." and you try to drag him up the wheel chair access ramp in front of the general practitioner. The access ramp has an incline of 23° . Your little brother has a mass of 43 kg and a static friction number μ_s on 0,12. What force must you apply to get your brother to start moving?

$$F_{push} = F_g(\sin\theta) + F_s = mg(\sin\theta) + \mu_s mg(\cos\theta)$$

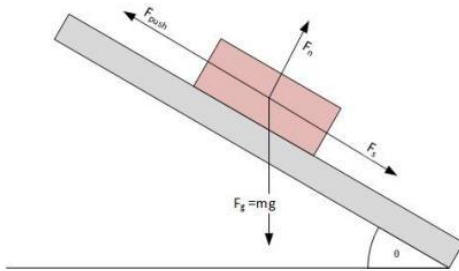


Figure 1 Assignment illustration for friction

Equations are shown in in the Excel spreadsheet and you need to convert equations to «something Excel understands». In addition, you need to read about sine and cosine functions in Excel in order to use them correct.

Answer alternatives:

211,47 N (the correct answer)

Superb. Keep your brain cool when entering long formulas. It's often smart to split up calculations in parts that give you a good overview of what is going on.

-50,61 N (wrong answer)

Several possibilities to tread wrong. Look up in help for how to use sine and cosine. Look up the RADIANS function together with sine and cosine. Otherwise, this is just a multiplication and summation of all factors. If you choose this answer you probably multiplied with PI() and forgot to divide by 180.

-383,93 N (wrong answer)

Here are several possibilities to tread wrong. Look up in help for how to use sine and cosine. Look up the RADIANS function together with sine and cosine. Otherwise, this is just a multiplication and summation of all factors. If you choose this answer, you probably did not use RADIANS.

4 ATOMIZED APPROACH

As mentioned, there are about 40 quizzes used as voluntary assignments. Each quiz has 3 to 5 questions with a minimum of three possible answers each – some with several more. In a review, we counted 590 possible answers - right and wrong. All these answers are the core in our atomized approach in that we provide feedback on all these answers. We have put a lot of effort in giving meaningful explanation for all wrong answers. The intention is to explain why this answer is wrong and how the student came to choose this answer. Included to direct students towards how to find the correct answer, without openly unveil the correct answer.

In effect, this gives the student a formative assessment for each test. The feedback is immediate and continuous and on demand. In addition to the formative assessment, each assignment also gets a summative assessment upon completion. Since instructors/TA do not use time to evaluate results on assignments, they can guide / tutor individual students who need follow up. The lab activities both in physical lab and the online meeting activities is spent on guiding / tutoring students instead.

Both formative and summative assessment are intended to encourage the student to keep working on voluntary assignments. Voluntary assignments prepare students for final exam and it is crucial since the exam uses the same mechanics as the voluntary assignments, though without formative assessment, only a summative, which gives the final grade.

4.1 Question quality

In the process, we discovered from time to time that it was difficult to make constructive feedback to possible answers. On many of these occasions we found that the question was not formulated good enough. Sometimes the question or task was ill formulated and reformulating was in place. Other times we found that the question all together was not good enough and was replaced. This indicated to us that more work was needed to assure better quality of questions. Each question had to serve two objectives:

- A problem based question (one that requires external tools, and not merely remembering phrases from instruction video)
- Plausible and meaningful answer alternatives with formative assessment on each alternative.

By doing this work, we not only provided formative feedback. But we also improved the questions themselves.

4.1.1 Question types

At present, we use the Canvas LMS. As of fall 2017 it provides the following question types:

- Multiple choice
- True / False
- Fill in the blank
- Fill in multiple blanks
- Multiple answers
- Multiple dropdown
- Matching
- Numerical answer
- Formula
- Essay
- File upload

Since we depend on question types that provide “correct answer” (automatic grading) some of the listed question types require intervention and we avoid the use of these. These are File upload and Essay type questions. Question types with only two outcomes, like True/False and Numerical answer, is ok in some cases, but will often give to limited possibilities for feedback. Other types like Matching and Formula have great potential, but can be limited by their implementation or their possibility for feedback in the actual LMS. Multiple choice is, by far, the most used in our assignments. The remaining question types are used to some extent.

4.2 Creating questions, and finding answer to them

The initial process of making questions was done before we started to incorporate formative feedback. Back then, this task was trivial and centered around the details we wanted the students to have focus on. In making the (wrong) alternatives it was a key aspect that the students could not merely guess the right answer, but had to complete the assignment to find it.

Making new questions with formative feedback requires more consideration.

4.3 Iterative development

One feasible approach to make questions with formative feedback on each possible answer is to use iterative development. One iteration results in a question, answer, feedback-set that can be used.

The process is both iterative and incremental. On each iteration the quality of the problem, possible answers and/or the feedback is improved.

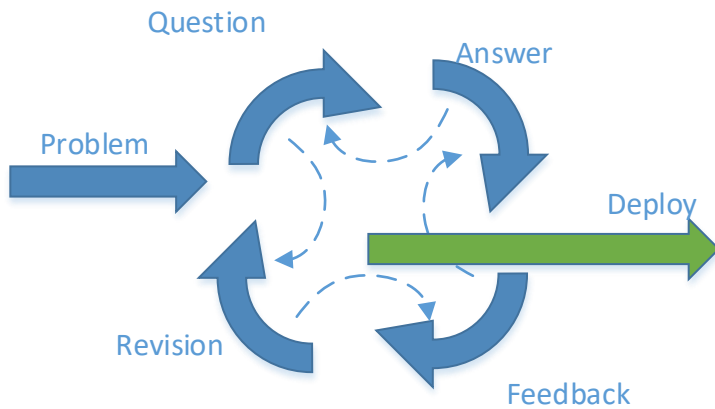


Figure 2 question development cycle

Each step: question, answers, feedback has its own loop of evaluating that can result in change in this or earlier step.

This approach works well for making new questions, adding feedback and for maintaining existing ones.

5 PRELIMINARY RESULTS

By comparing student grade against how many assignments the student completed, we found a correlation between the two. Not surprisingly, the more assignments students completed, the better they performed på the final exam and the fail percent on final exam decreased. In 2014 the fail percent on final exam was 18 percent. We speculate that part of a high fail percent was in part due to other factors such as initial startup load. In 2015, we did some small adjustment in the administrative region of the course, which contributed to better organizing. Better organizing and no start-up challenges reduced the fail percent to 9,69 percent. In 2016 we started experimenting with atomized feedback and saw that students did more work with assignments and we believe this contributed to give a pronounced drop in exam failure down to 3,98 %

Our motivation to make changes to the pedagogical method partly emerged from findings in Figure 3 to Figure 5. We discovered that several students did all the assignments but still failed on the final exam. This information are to the left side in area marked "A" in Figure 3 to Figure 5.

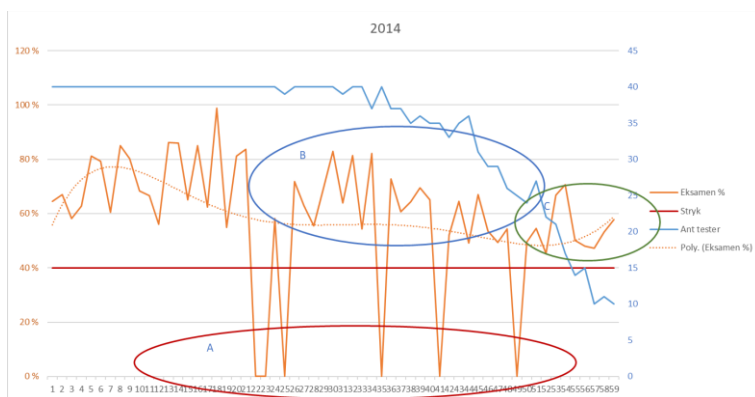


Figure 3 Number of assignments versus grade 2014

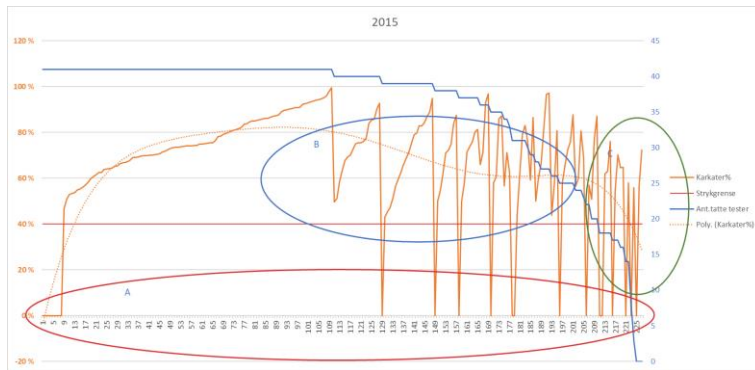


Figure 4 Number of assignments versus grade 2015

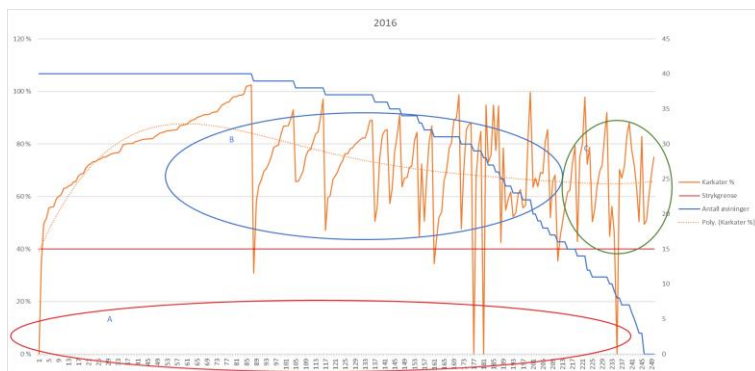


Figure 5 Number of assignments versus grade 2016

We did not dive into and questioned these individual students, and ask them how and why they worked so hard but still failed the final exam. We acknowledge that students could do all assignments (they could get help) and still not “have done their homework”. However, conferring a few random students with this issue, two related issues emerged: Several questions on the final exam had too much text, it took too much time to read, i.e. get to the point, one doesn’t need to paint the whole scene for a simple question. The other was they did not understand (did not get feedback) what they had done - they were guessing. The guessing issue led to insight that we needed more feedback to students about what the questions was about and why and how to solve them.

After implementing atomized formative assessment, we have noticed a significant reduction in enquiries about assignments and instructions. Student assistants are approached less, teacher and staff are approached less with questions about assignments. i.e. the course stands on its own feet.

It’s too early to state whether the average final exam grade has moved from today’s “C” since the evaluation is not finished at time of writing. However, there is a week trend from 2014 to 2017 that the average gauss distribution is sliding towards an average “B”

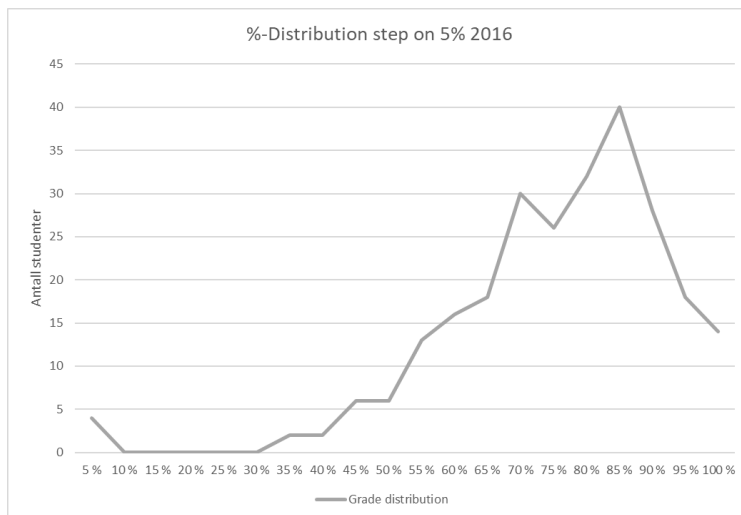


Figure 6 Grade distribution

6 FURTHER WORK

The quiz engine of Canvas has several known minor faults and a new quiz engine is in the beta test phase. Our hope is that this engine is released soon so we can take advantage of several proposed new features

Canvas has the ability to log a wide range of parameters. We are interested in data about how much time is spent on working on each assignment. Time spent on assignment gives indication on degree of difficulties. This would be an important factor in adjusting the instruction and the assignments accordingly.

Next step in this process would be to use artificial intelligence to predict performance and use this to generate individual feedback.

7 REFERENCE

- [1] R. E. Bennett, "Formative assessment: a critical review," *Assessment in Education: Principles, Policy & Practice*, vol. 18, pp. 5-25, 2011/02/01 2011.
- [2] D. J. Nicol and D. Macfarlane-Dick, "Formative assessment and self-regulated learning: a model and seven principles of good feedback practice," *Studies in Higher Education*, vol. 31, pp. 199-218, 2006/04/01 2006.
- [3] B. A. Bertheussen. (2014). *Automatisk formativ feedback kan gi god motivasjon og læring*. Available: <https://utdanningsforskning.no/artikler/automatisk-formativ-feedback/>
- [4] K. Bjørkli and K. Arnesen, "Kombinert summativ og formativ vurdering i matematikk," *Uniped*, vol. 38, pp. 337-344, 2015.
- [5] N. Frey and D. Fisher, *The Formative Assessment Action Plan : Practical Steps to More Successful Teaching and Learning*. Alexandria: ASCD, 2011.
- [6] J. Hattie and H. Timperley, "The Power of Feedback," *Review of Educational Research*, vol. 77, pp. 81-112, 2007.