How we can use online feedback to maximise engagement with the assessment criteria?

UCL E-Learning Development Grant project report, August 2016

Dr Elodie Douarin, Lecturer in Economics, School of Slavonic and Eastern European Studies, UCL.
Dr Mira Vogel, Digital Education Advisor (SLASH), UCL.

"I only have hopes, so it will be obvious for the teacher to use our recording and to understand what we'd like to change to deeply understand the grades. And uh … thank you very much for your time."

Helen - Topics in Microeconomics student

Contents

Contents

Executive summary

Overview

In 2015 we received an E-Learning Development Grant of £2,000 from UCL Digital Education to investigate how digital feedback could maximise engagement with assessment criteria.

Our context was the Topics in Microeconomics module, taken by 96 2nd year undergraduates from three different programmes. The module was assessed through a final exam (worth 75% of the final mark) and a coursework (25% of the final mark). Both the exam and the four-part coursework included essay writing, calculations, and production of graphs. They were preceded by an optional formative assessment which required students to solve half of a past-exam paper. This study explored the students' experiences as they undertook the formative assessment and summative coursework, focusing on their use of the marking criteria and their perceptions of the feedback they received. They engaged with marking criteria and feedback first in Moodle Assignment and second in Turnitin Assignment.

Methods

To investigate how students used the marking criteria and feedback and how they encountered the feedback they received on this course, we conducted an initial voluntary survey of their views on assessment criteria and feedback (response rate: 52%; number of respondents: 50). We explored the relationship between these views and other information drawn from data routinely and automatically collected within UCL, including past marks, performance on Topics in Microeconomics, attendance and information collected through Students Evaluation Questionnaires. A self-selected sample of 9 students participated in more depth by recording their responses to a set of questions at the time they encountered their formative and summative feedback. They also took part in a focus group discussing their experience with feedback on Topics in Microeconomics. Transcripts were analysed using NVivo.

Findings

The survey illustrated clearly that **students engage little with the marking criteria and find it difficult to use them to improve their performances**. It also showed that the students who did use the criteria did not necessarily find it easier to judge the quality of their own work or to understand the feedback or mark they had received.

Using information about assessment performance in Topics in Microeconomics, we show that the students who submitted the voluntary formative assessment did better in their summative coursework. **Their marks are on average 4 points above the marks of those who did not submit their formative coursework**. This difference is statistically significant and robust to the addition of a number of controls such as performances in Introduction to Microeconomics (a compulsory pre-requisite for Topics in Microeconomics), attendance, and a measure of conscientiousness (tutor assessed efforts in tutorials). Thus **this demonstrates the benefits of engaging with the marking criteria and receiving feedback on their work**.

In the Student Evaluation Questionnaire for the course, students also reported to be satisfied overall with Topics in Microeconomics and for **all questions relating to assessment and feedback in particular, they reported positive views on average**. However, the analysis of the information collected through feedback walk-through and the final focus group **shows that students' expectations about feedback had <u>not</u> been met**.

The recordings and focus group data allowed us to summarise what participants considered to be helpful and unhelpful, along with any remedy they proposed.

Student participants' views of assessment criteria		
Helpful	**Unhelpful**	**Notes**
Related to the marks – specifically where they lost marks.	Criteria which are "abstract" or "generic".	Community measures to build shared understandings may persuade students to come to terms with the nuance in criteria. Moodle and Turnitin to some extent, and differently, support giving feedback for each criterion.
Each question has its own marking grid.	As above.	Although this may be helpful as feedback, it is unlikely to improve criteria validity. Moodle and Turnitin can enable this.
Make the weighting for each criterion explicit.	As above.	Students asked for better guidance about where to spend their effort and word limit. Moodle & Turnitin can enable this.
Criteria which were conspicuously located.	-	In the case of less conspicuous criteria, some students were entirely unaware they existed. Moodle is better at making criteria, and feedback given via the rubric, salient.
Criteria which are referenced in the feedback.	Criteria which markers don't use.	There was a widespread assumption that markers don't refer to the criteria, and little concept of the criteria as educationally helpful (beyond just grading).
Model answers in sufficient detail to illustrate how the criteria can be met.	-	Some students requested exemplar responses which illustrated what good and less good work looks like.

Table 1- participants' views of what makes criteria helpful or unhelpful

Student participants' view of assessment feedback		
Helpful	**Unhelpful**	**Notes**
Explanation of mistakes.	Cryptic feedback consisting of questions, or just one or two words, or ticks and crosses.	The tutor hoped students would refer to the model answer. However, this was not clear to the students. Where ticks and crosses are used, as well as requiring further explanation, label lines may be needed to clarify what they refer to.
Specific, constructive advice.	Pointing out problems without giving advice or direction.	Feedback which students feel confident they can transfer into improved marks.
Feedback which explained where marks had been lost or could be gained.	Feedback which didn't explicitly relate to the mark.	Moodle and Turnitin both to some extent, and differently, enable tutors to associate marks with criteria (which constitutes feedback).
Feedback which related to the criteria.	Feedback which didn't relate to the criteria.	Moodle and Turnitin both to some extent, and differently, enable tutors to associate feedback with criteria.
Feedback commensurate with the marks.	Feedback at odds with the marks.	Positive comments relating to 'surface' aspects which do not carry many marks, or which are intended as encouraging, need to be explained to students.
Timely feedback	Late feedback.	As time passes students struggle to remember their thinking when they were undertaking the assessment.
Digital feedback	Handwriting.	Digital feedback is legible and readily accessible.
-	A lot of feedback dispersed across several modes.	This greatly increased the effort students had to make to relate each piece of feedback to others, and to their own work. Not all students undertook the effort. Moreover they were liable to miss an aspect of feedback, indicating that feedback dispersed between different modes needs to be obviously signposted.

Table 2 - participants' views of what makes feedback helpful or unhelpful

An important point coming out of this exercise is that, particularly on Turnitin, students could not find some elements of feedback, and on both platforms they often struggled to relate the comments received to their work and the solutions. This is partly for technical reasons (which have since been resolved) but it is also due to the multiple modes of feedback available to students for each piece of work. Triangulating the different modes of feedback proved challenging for students, as they sometimes had to refer to their original work, a separate document with the solution, comments embedded in the text of their work, the rubric, and general comments at the end or on the side of their paper, to identify their errors.

Regarding the form of feedback given, specific, constructive advice and explanations of errors were well received. However, positive comments on successful elements of assessment, or prompts encouraging the students to look for the correct answers, were sometimes viewed as vague, unclear or unhelpful. It appeared that for exercises requiring numerical and graphical answers, students were keen to be given model answers (exemplars) and to be told explicitly where they had lost each mark and why.

The focus group also illustrated a low level of trust that marking criteria were actually used in marking. This may stem from a lack of shared interpretation of the criteria among staff and students, or unfamiliarity with the criteria preventing students from recognising when markers were in fact using them. With this in mind, it is important to note that even when prompted, students did not discuss the criteria descriptors in any detail. It is not clear whether this indicates a lack of familiarity with these descriptors, or alternatively that students were sufficiently confident in their interpretations that they felt no need to discuss them further.

In the final ‘Key findings’ section we summarise a set of principles derived from the literature, student experiences and tutor experiences described in this report. We hope this will be a solid basis for future practice with criteria and feedback.

End of Executive Summary

Background

Assessment criteria

In this study we presented the assessment criteria in digital format as rubrics (see 'Appendix 4 – assessment criteria'). In their review of rubric use in higher education, Reddy and Andrade (2010) define rubrics as:

> "… a document that articulates the expectations for an assignment by listing the criteria or what counts, and describing levels of quality from excellent to poor."

They observe rubrics to have three broad functions. The most obvious is a standard by which to grade students' work. Although rubrics used in this way can build students' confidence that the grading is fair and transparent, the authors found contrasting attitudes among tutors, ranging from positivity to resistance. They attributed the resistance to a limited conception of rubrics as exclusively for assigning grades rather than as an educational opportunity for students,

> "… to engage in important processes, including identifying critical issues in an assignment and, thereby, reducing uncertainty and doing more meaningful work, determining the amount of effort needed for an assignment, evaluating their own performances in order to get immediate feedback, especially on weaknesses, estimating their grades prior to the submission of assignments and focussing their efforts so as to improve performance on subsequent assignments."

However, to gain the benefits of assessment criteria, students need to have access to the rubric at the time the assignment is given, enabling them to refer to it before the deadline. Moreover, availability alone is insufficient to improve educational outcomes - students need to be encouraged and guided to "engage deeply with rubrics, perhaps by co-creating them and using them for self- and peer assessments". This is in line with the conclusions of Bloxham and colleagues (2015) who argued that criteria cannot bring about consistent marking. They were clear that this should not be taken as a criticism of markers (c.f. Bloxham and colleagues, 2011) and also argued against more detailed assessment criteria, since these might both interfere with rewarding unexpected responses, and also encourage staff to work "backwards from holistic judgements to determine commensurate marks for individual criteria". Instead they emphasised nuance of interpretation and the importance of "community processes aimed at developing a shared understanding of assessment standards".

A third potential of rubrics, highlighted by Jisc (Ferrell, 2013) is as evidence for a teaching team or department. The aggregated data from rubrics can illuminate attainment of the knowledge, skills and attributes which are taught and learned at a cohort level, and inform improvements at a module or programme level (or higher).

Assessment feedback

Action without feedback is unproductive for learning. A numeric mark is a consequence of the work students submit, whereas feedback is a deliberate and purposeful comment on it. Sadler's definition of formative feedback is helpful here: “to shape and improve the students’ competence by short-circuiting the randomness and inefficiency of trial and error learning” (Sadler, 1989, p. 120).

For educators, giving assessment feedback entails discerning the potential of a student or group, and making well-judged interventions which are not mechanistically mistaken by students for solutions, but build on each other to help students attain their most potentially advanced state for that particular stage of learning or - more concretely - attain the reference level of knowledge specified by the assessment criteria. Orsmond and colleagues (2005) identify this scaffolding and bridging as a particularly demanding practice for educators.

Higgins and colleagues (2002) investigated how 19 students (in this case, from Business and Humanities) responded to their tutors' comments – i.e. how feedback works – and found that students encounter problems when trying to use feedback:

> “...students in our study perceive feedback negatively if it does not provide enough information to be helpful, if it is too impersonal, and if it is too general and vague to be of any formative use. Handwriting also seems to be a common problem. For example, 40% of our questionnaire respondents often found feedback comments difficult to read.”

Another barrier to use of feedback, students may not share common terms (for example, the assessment criteria) with their tutors:

> “Often, feedback comments employ the academic language used to express assessment criteria, but only 33% of our respondents claimed to understand these criteria. An inability to fully comprehend the meaning of assessment feedback may not necessarily prevent students from paying attention to tutors’ comments, since they may unknowingly interpret them incorrectly yet still attempt to utilise them. Nevertheless, this will almost certainly present an obstacle for many.”

These doubts about how or even whether students use feedback amount to a disincentive for tutors; if tutors perceive that students are not making good use of their feedback, they may be less inclined to take pains producing it. Crisp (2007) asks whether feedback is worth the effort, quoting Pardoe (2000):

> “Reading and marking texts that students have written can be a dispiriting experience. Sometimes it seems that key points of the course have not been registered in the students' minds. Often the texts seem even to lack a basic knowledge of writing 'that surely should have been acquired long ago'.”

Yet, as Higgins and colleagues (2002) point out, students never seem to lose faith in the potential value of feedback, and consider it something they deserve.

Investigating six different forms of feedback in six focus groups, Lipnevich and colleagues (2009) reported a high level of agreement among students about the ideal qualities of feedback:

> "The most salient finding of the experiment was that descriptive feedback specific to individual work is critical to students' improvement. Detailed, specific, descriptive feedback which focuses students' attention on their work, rather than the self, is the most advantageous kind of information for students. The benefits of feedback occur at all levels of performance. Evaluative feedback in the form of grades may be helpful if no other options are available, and can beneficially be accompanied by some form of encouragement."

Black and Wiliam (2009) summarise the *joint* responsibility students and teachers have for learning:

> "… since the responsibility for learning rests with *both* the teacher *and* the learner, it is incumbent on each to do all they can to mitigate the impact of any failures of the other (in the language of partnership law, teachers and learners are jointly and severally liable!)."

In the 2016 National Student Survey results for the questions on Assessment and Feedback (questions 5-8), the overall score for UCL SSEES Economics was 66, a sharp improvement on 2015 but still the lowest-scoring area of the NSS (for comparison purposes, the UCL average for Assessment and Feedback was 64, the sector average was 74. The Economics Network (2016) shows NSS trends for Economics and related subjects across the UK.

Electronic management of assessment

In her 2016 survey of the electronic management of assessment (EMA) scene for the UK higher education Heads of E-Learning Forum, Barbara Newland found that although a minority of institutions have policies on e-feedback or e-return, both have increased steeply in recent years. In 2016 e-feedback is the only form of feedback for over 28% of institutions. Institutional policy on e-marking has seen the least change but has still roughly doubled since 2013 to around 28% in 2016. The proportion of institutions considering or working towards an integrated EMA approach is growing, though not very fast. The clue to this slow growth may be found in the responses to Newland's questions about integration: records management and automation of routine aspects have only been achieved by a small minority of institutions. The proportions of staff with negative feelings about e-marking and e-feedback have fallen slightly, but for e-marking remains at 18%. It is interesting to note that academics report doing more work to set up e-assessment processes than administrators report.

As part of the Jisc Assessment and Feedback Programme (2011-14), 40 UK high education providers reviewed and enhanced their practice with appropriate use of technologies. The synthesis report (Ferrell, 2013) observed that "online marking has reached a level of maturity whereby the evidence for its effectiveness and efficiency may soon result in a critical mass of institutions and academics adopting the practice. This study seeks to contribute to this evidence base.

Students of Topics in Microeconomics

The study focused on Topics in Microeconomics (SESS2005), a second year module which is usually taken by all students on the BA EBEES (Economics and Business with East European Studies) programme at SSEES, it is also opened to BASc students on the Societies pathway and to other students within SSEES (students on the Politics with East European Studies or PEES programme for example). All students on the course have studied through Introduction to Microeconomics (SESS1003), it is a compulsory pre-requisite.

Topics in Microeconomics is an intermediate-level microeconomics course focusing on household decisions (consumption, investment and labour supply), it also includes an introduction to welfare economics, auctions and game theory. The course is quite dense as it covers many different topics. It is also a little bit technical: students need to be comfortable with differentiation and able to draw and analyse functions. It is delivered through a weekly 2 hour lecture and a 1 hour seminar (typical size: 17 students). In the past the average mark for the module has been around 60% but often with large standard deviations, and with failure rates varying year on year but generally above 10%.

There is quite a lot of heterogeneity in the typical Topics in Microeconomics cohort. Students on the EBEES programme typically take a lot of economics and business related courses and are therefore significantly more exposed to economics concepts and thinking than students on the PEES or BASc programme. Students on the EBEES and BASc programmes tend to have more opportunities to take quantitative courses than students on the PEES programme. The entry requirement for EBEES includes an A-level in Maths, but not for PEES or BASc. Some students on all programmes have taken economics for their A-level and may be more comfortable with Economics intuitions than those who have not.

The lecture is delivered in a relatively traditional way. Key concepts are presented and explained by the lecturer, at regular intervals students are offered an opportunity to work through simple exercises or graphs, or to discuss a concept with their neighbour. An effort is made to link the content covered to other courses the students may follow (including in other disciplines) and to discuss the limitations of the models presented. A well-regarded textbook is used as an alternative source of explanation, and as an exercise bank.

The seminar occurs weekly, after the lecture. Each seminar focuses on solving 3 to 5 exercises relating to the weekly lecture. Students are strongly encouraged to attempt to solve the problems by themselves ahead of the seminars, and to actively participate in seminars by solving problems at the board or asking questions. In previous years, students have been given time to compare their answers with peers at the beginning of the session, to build their confidence and facilitate active participation. However, timetabling issues have sometimes meant that seminars were scheduled too soon after a lecture to allow students to prepare for the seminars. Seminar tutors have also sometimes reported feeling pressured to provide answers to students rather than encourage them to solve the problems themselves, due to time constraint and to students' resistance to participating.

In 2016, there were 96 students registered on Topics in Microeconomics, 6 were from PEES, 14 from BASc (14.5%) and the rest (76) from EBEES. Women accounted for 50 percent of the class.

The assessment for Topics in Microeconomics has three elements:

1. A voluntary element of formative assessment: Early in the term, students are offered to solve a subset of questions from a past exam paper. Comprehensive feedback is provided on the work submitted and a mark is given
2. A compulsory element of formative and summative assessment, worth 25% of the final mark: towards the end of term students are asked to solve a set of exercises, typically more challenging than the exercises found in the exam paper
3. A compulsory end-of-year exam, worth 75% of the final mark for the course.

All three elements of assessment include exercises that will be solved using calculations or graphical analyses. Students are asked to explain their approach in their own words in all instances, as mathematical formulas or graphs alone are insufficient to demonstrate one's understanding, but at least 1 question in each element of assessment will explicitly ask for an essay-type answer. In 2016, each piece of assessment was submitted individually (no group work).

For each element of assessment, students are invited to refer to the module "assessment criteria" to understand how their work is assessed. These assessment criteria are presented in 'Appendix 4 – assessment criteria', and are adapted from the generic set of assessment criteria that are used at SSEES.

Methods

All work with students was carried out in line with UCL's ethics procedures. Informed written consent was obtained and students were free to end participation at any time. Due to the potentially sensitive nature of the contributions and the liability this posed to frank responses, we decided to take measures to reassure students that they could not be identified from their responses. Students responded to the questionnaire, uploaded their recordings, completed the questionnaire, and signed up for the focus group in a Moodle space from which SSEES staff were excluded. The nine students who participated in depth were given aliases and had the opportunity to validate their transcripts and this report. They could also check that their contribution was not identifiable. Where this was in doubt they were invited to propose changes.

At the beginning of the module all students were asked to complete an **anonymous questionnaire** about a previous experience with assessment and feedback at UCL. Questions were developed from conversations with students who had taken Topics in Microeconomics in 2015/16, and refined and validated with Student Academic Representatives from the current cohort. For the questions, see 'Appendix 1 – questionnaire'. Completion was incentivised by promising the students a set of answers for a past exam paper if at least 75% of the cohort filled the questionnaire. The questionnaire remained available online for three weeks and the final response rate was 52% (the promised past-exam solutions were distributed despite this target response rate shortfall).

Additional information typically collected on UCL courses was analysed to complement the questionnaire data. First, a short analysis of the marks of students on the course, using past performances, attendance and participation in seminars is presented to illustrate the benefits of taking

part in voluntary formative assessment. Second, information from the Student Evaluation Questionnaire are also briefly reported.

Early in the module a self-selected sample was recruited to contribute in more depth during the module. We aimed for 20 students, but recruited nine. We gained their informed consent to participate according to '**Error! Reference source not found.**'. Each of the nine students contributed to two or more of these activities. Participation was incentivised at £14.32 per hour they spent on the project.

Participants completed a **formative coursework in Moodle Assignment** followed by a **summative coursework in Turnitin Assignment**. In each instance the assessment criteria were given as a **rubric**, and **feedback was given in four modes**:

- contextualized 'bubble' comments in the submitted text;
- an overall comment;
- an indication of achievement for each criterion, indicated by a shading of those rubric levels.
- a solution or 'answer key' for each section – an unelaborated response (briefer than expected of students) and some pointers about common mistakes.

The assessor also gave students a numeric mark.

The release of feedback in relation to assessment deadlines gave students the opportunity to use their formative feedback in their summative coursework, and their summative feedback in their examination.

On the two separate occasions they received marks and feedback for their formative and summative coursework, **students recorded their responses** to a set of questions (see 'Appendix 2 – questions for students' recordings') using screencasting software, or audio recording software, or a word processor. Again, Student Academic Representatives were involved in developing the questions. Support for using open source Windows, Mac and Linux compatible recording software was available in the form of a step-by-step video walkthrough. Responses were transcribed.

In revision week before the examination, the nine students mentioned above participated in a **focus group** (see 'Appendix 3 – focus group questions'). To improve participation, this was piggy-backed on a revision session and lunch was provided. Again, Student Academic Representatives were involved in developing the questions. The focus group was recorded and transcribed, and participants validated and checked the transcript before we analysed it.

The **tutor**, as a co-investigator and marker, fed her experiences into the project by email to the Digital Education co-investigator, in person, and as she authored this report. We have decided to integrate the experiences of students and tutor-marker since policy and practice with assessment require that staff and student needs be considered and balanced holistically.

All of the transcriptions were analysed together along with any word-processed data students had contributed. We used Nvivo to open-code the responses by theme. Because of the need to ensure that students could not be anecdotally identified from their responses, this part of the analysis was carried out solely by the Digital Education investigator.

Finally, participants were invited to **validate the final report** before it was published more widely. Of the four who fed back, one proposed specific amendments around emphasising the importance of model answers and highlighting the low marks received for one part of the assessment – these were incorporated. Two others were very positive about the report and one had nothing further to add.

Activity	Time / minutes	Students	Data
Questionnaire	5	50	Responses to 5 questions, plus comments.
Recording themselves accessing formative assessment feedback	60	7	90 minutes of recorded data (4 video, 2 audio) and 1 document.
Recording themselves accessing summative assessment feedback	60	7	55 minutes of recorded data (3 video, 2 audio) and 2 documents.
Focus group	90	9	76 minute audio recording.
Validating focus group transcript	30	4	N/A
Validating final report	60	4	N/A

Table 3 - participant activities

Questionnaire analysis

For questionnaire questions in full, refer to 'Appendix 1 – questionnaire'.

The respondents profile was roughly representative of the cohort, with 22% of the respondents being BASc students (compared to 14.5% of the cohort being from the BASc) and with 58% of the respondents being female (compared to 50% of the cohort). Slightly over half the cohort did fill the questionnaire (50 respondents). Despite failing to reach the target 75% response rate, the incentive was released to all students.

The responses are summarised below.

Question 1 - Thinking about the assessment criteria for your last piece of coursework (for any module), which of the following statements do you most agree with?

62% of the respondents said they had difficulties understanding the criteria. But even when the criteria were thought to be clear, nearly all students found it difficult to relate them to their own work.

Question 1	Share of respondents
1. I was not aware of the criteria	26%
2. I was only vaguely aware of the criteria	20%
3. I had difficulty understanding the criteria and difficulty relating them to my own work	16%
4. The criteria seemed clear but in practice I had difficulty relating them to my own work.	36%
5. The criteria were clear, and I was able to relate them to my own work	2%

Question 2 - Still thinking about your last piece of coursework, which of the following best describes how you used the assessment criteria?

46% of respondents said they did not consider the criteria at all.

Question 2	Share of respondents
1. I did not consider the criteria	46%
2. I considered the criteria at the start of preparing my work.	22%
3. I considered the criteria as the deadline approached.	12%
4. I considered the criteria throughout.	12%
5. I systematically tried to apply the criteria to different aspects of my work throughout.	8%

Question 3 - Still thinking about that last piece of coursework, how well do the following qualities describe the feedback you received?

Students were asked whether the last piece of feedback they had received was informative, easy to understand, helpful, related to the assessment criteria, related to their work and timely.

Two third of the respondents stated that their feedback was timely (agree + strongly agree: 68%) and four fifths that it related to their own work (agree + strongly agree: 76%). These are the two areas were student's assessment was the most positive. At the other end of the spectrum, only a third of the respondents stated that their feedback related to the assessment criteria (agree + strongly agree: 37%), and less than half of the respondents found their feedback helpful (agree + strongly agree: 43%). Just over half found their feedback informative (agree + strongly agree: 53%) or easy to understand (agree + strongly agree: 57%).

Question 3	Share of respondents					
	Informative	Easy to understand	Helpful	Related to assessment criteria	Related to my own work	Timely
1. Strongly disagree	8%	6%	18%	13%	2%	12%
2. Disagree	29%	21%	25%	17%	6%	10%
3. Neither	10%	16%	14%	33%	16%	10%
4. Agree	**51%**	**49%**	**37%**	**33%**	**68%**	**56%**
5. Strongly agree	2%	8%	6%	4%	8%	12%

Question 4 - How did the mark you received for that coursework compare to your expectations?

38% of students received a mark in line with their expectations. The others either did not know what to expect or were surprised by their marks. Students are more likely to over-estimate than under-estimate their marks, a priori.

Question 4	Share of respondents
1. Higher than expected	18%
2. Roughly the same	38%
3. Lower than expected	38%
4. I did not know what to expect	6%

Using the answers above, we can perform some simple regressions, to identify how different answers correlate with others.

First, we took students' responses to each question listed above and regressed it against their gender and degree programme. This highlighted that BASc students find the assessment criteria significantly clearer than the rest of the sample, and they also use them significantly more when working through an assignment. Indeed, the average answer to Questions 1 and 2, is 1.1 and 1.4 points higher respectively for BASc students compared to the rest of the respondents. This is probably a result of efforts on the part of BASc staff to publish assessment criteria on the moodle page of every BASc core course and to mark work using rubrics clearly referencing the criteria. However, it is interesting to note that BASc students do not find feedback more useful than the rest of the cohort. Looking at Question 3, the BASc students always reported a lower score, showing that students from the BASc rate the feedback they have received less positively than their peers, however the difference is only significant in one instance: when answers to Question 3 were analysed (regarding how helpful the feedback received was). It is important to note however that we do not know which module the students were referring to in their answers, so BASc students may not be commenting on BASc modules.

Gender is never significantly correlated to any of the responses provided, so there are no clear differences in views between male and female respondents.

Second, we performed simple regressions, to see whether different statements were correlated. There we found that a higher rating for Question 1 (awareness of criteria) and 2 (whether they were used) tend to be positively correlated with receiving a mark in line with expectations and negatively correlated with receiving a lower mark than expected. So it seems that those who do understand and use the marking criteria may be slightly better at assessing the quality of their work than those who do not. Not using or understanding the criteria seems to be very weakly associated with an over-estimation of one's performance. This would be in line with expectations, however these associations are not statistically significant.

Overall this small survey provides a sound motivation for the study which followed. It shows that students have a limited understanding of what the marking criteria mean, and rarely use them. It also shows that making use of the criteria is not necessarily associated with performing better, assessing the quality of one's work more accurately, nor rating the feedback received as better. These findings are in line with the literature on criteria and feedback.

Benefits of submitting a voluntary formative coursework

To further contextualise the focus group and students' views of feedback they had received, we propose a second quantitative analysis. Focusing on students' achievements, we want to assess whether students have benefited from submitting a piece of coursework before their final exam. The implicit question we are asking is whether students appear to have learned from the feedback received or whether they have been able to build on this feedback to improve their performance in the next round of assessment. To do so, we have again run some simple regressions aimed at relating students' achievement in one round of assessment to previous performances. The models are reported in "Appendix 5 – Quantitative Analyses". The key findings are reported below in non-technical terms.

Before going into the regressions, we ought to describe the performance of the students on the course. As stated earlier, Topics in Microeconomics is an intermediate level course offered to second year students. All students on the course have taken Introduction to Microeconomics in their first year at UCL, as this is a compulsory pre-requisite. We can therefore take the student's mark in Introduction to Microeconomics to reflect their entry-level knowledge of Microeconomics. We measure attendance (number of tutorial attended out of 10) and conscientiousness (as perceived by the class tutor based on the student's attitude in seminar, measured on a scale of 1 to 7 with 1 reflecting minimal involvement and efforts, and 7 standing for outstanding conscientiousness)[1]. Due to missing data, this section of the report is based on information about 89 students only (i.e. 7 students from the cohort did not enter the analysis).

[1] The class tutor in charge of all the tutorial groups was provided with the following definition of conscientiousness:

> "Conscientiousness is the personality trait of being thorough, careful, or vigilant. Conscientiousness implies a desire to do a task well. Conscientious people are efficient and organized as opposed to easy-going and disorderly" (Wikipedia).

She was then asked to give a score to each student on the course, using a 7-point scale where point 4 is 'average efforts or conscientiousness' and each point above 4 indicate an incremental increase in conscientiousness up to 7 which should be given to the most conscientious student, and similarly each point below 4 indicates an incremental decrease down to 1 which should be given to completely passive students. To avoid any misunderstanding, the following clarification was later provided:

> 'Conscientiousness for Topics in Micro should translate into coming ready to seminars, being willing to ask questions (maybe after the session for more introvert students) or to demonstrate their results at the board. But it is not about getting the right answer'.

	Variables	Total sample N=89		Sub-sample of students who submitted a formative assessment N=55 61.8% of 89	
		average	**range**	**average**	**range**
Outcomes: Marks obtained for Topics in Microeconomics	**Exam (mark)**	63.0	[32 ,85]	64.6	[32 ,85]
	Coursework (mark)	62.5	[23 ,85]	64.1	[23 ,85]
	Formative (mark)	n/a	n/a	45.9	[0 ,76]
Controls	**BASc**	0.135	[0 ,1]	0.163	[0 ,1]
	Intro to Micro (mark)	66.4	[40 ,85]	67.3	[47 ,85]
	Attend	8.6	[6, 10]	8.7	[7, 10]
	Conscientious	4.9	[2, 7]	5.0	[2, 7]
Difference in marks Coursework - Formative		n/a	n/a	18.2	[-25, 58]

Table 4 - Description of sample

From the table above, we can see that the average mark for the end of year exam was 63, while the average mark for the summative coursework was 62.5. These figures are in line with what is generally observed for second year courses at UCL and with the cohort averages observed in previous years for Topics in Microeconomics (the average mark for the module was 58 in 2014/15 and 63 in 2013/14). The take up rate for the voluntary formative assessment was a respectable 62%, but among the sub-set of students who submitted this piece of work, the average mark was a disappointing 45.9. However, the students who participated in the focus group indicated that they had spent little effort on their formative coursework because it did not count towards their final mark.

Looking at the share of students on the BASc programme, the marks received on average for Introduction to Microeconomics, the number of seminars attended and our 'Conscientiousness score', we can see that there are little differences between the total sample (13.5%, 66.4, 8.6 out of 10, and 4.9 out 7 respectively) and the sample of students who did submit the formative assessment (16.3%, 67.3, 8.7 out of 10, and 5 out of 7). The sample of students who did submit a formative coursework therefore seems representative of the total cohort in those dimensions. To confirm this, we regressed the decision to take part in the formative assessment against these 4 variables, and we found that all were positively associated with participation, except for conscientiousness, but only the relationship between attendance and participation was statically significant, albeit at a 10% level only. In other words, we find weak evidence that students with better attendance records are more likely to submit their formative coursework, while more able students (i.e. with higher marks in Introduction to Microeconomics), more conscientious students or students on different programmes are not likely to behave differently than the rest of the cohort. To some extent this is surprising as better students and more conscientious students might be expected to be more likely to engage in voluntary learning activities.

Next, we looked into the determinants of students' final marks. In other words, we regressed students' marks for their two summative assessments against their past performances, attendance, conscientiousness and degree, as well as their decision whether or not to submit the formative assessment. In doing so, we wanted to assess whether students benefited from submitting this voluntary piece of work. We found that indeed, students who did submit a formative assessment performed better in their coursework and exam than the rest of the cohort, even after controlling for their attendance, conscientiousness and entry level in microeconomics. Using interaction terms, we further found that students who had performed less well in their first year module, benefited more from submitting a piece of formative work.

This therefore demonstrates that engaging in voluntary learning activities has clear benefits. The increase in coursework marks observed for students who submitted a formative assignment could be explained by:

1. A benefit from engaging with the task and gaining a better understanding of the course;
2. A benefit from receiving feedback on the task and gaining a better understanding of the marking criteria.

Because the solutions to the formative coursework were distributed to the whole cohort, all students were able to work through the formative assignment in preparation for their coursework and exam. So this may slightly weaken the likelihood of explanation 1.

Further to this, and because the summative coursework was submitted before the final exam but after the formative coursework, students should have been able to gain further benefits from receiving feedback on their summative coursework, which could translate into better exam marks. However, one could also expect that this second coursework submitted before the exam would level off the benefits of submitting a formative coursework. But we find no evidence of this. The relationship between coursework and exam mark is not different for students who did or did not submit their formative coursework. In other words, whatever those who submitted their formative coursework gained from the experience is not levelled off once everyone has been through the compulsory coursework.

Overall, this shows that there can be a substantial benefit from submitting the voluntary coursework (here an increase of 4 marks for the next piece of summative assessment).

Analysis of the Student Evaluation Questionnaire

Only 12 students filled the standard ‘Student Evaluation Questionnaire’ at the end of the academic year. This low response rate (13%) can be explained by the switch to online SEQ, which may have made it easier for students to ignore them. We are still reporting the responses here as they provide additional information on students’ views of the module and the way it is assessed in particular.

First, on average these students rated the Topics in Microeconomics as good overall (Overall satisfaction with the lectures: 4.1/5 and with the seminars: 4.6/5). Regarding assessment and feedback, three questions were asked and the answers provided are reported in the table below. They show that the

level of satisfaction regarding assessment and feedback is in line with the overall satisfaction for the module (with average scores ranging from 4.0 to 4.5).

Statement	1 Strongly disagree	2 disagree	3 neither	4 agree	5 Strongly agree	Average score
The criteria used for marking and assessment were clear in advance	0%	8%	8%	33%	50%	4.2
The feedback provided on my coursework was helpful in developing my understanding of the subject	10%	0%	10%	40%	40%	4.0
I had the opportunity to receive feedback on a coursework prior to submitting an assessed work counting towards the final assessment	0%	0%	0%	55%	45%	4.5

To conclude, we note that students did benefit from submitting a piece of formative coursework, that the students who had submitted this formative piece are not necessarily better or more zealous, and that those who benefit the most are in fact those who did not achieve very high marks in their first year. We also note that students were overall satisfied with the course and the feedback received on Topics in Microeconomics. Overall this makes Topics in Microeconomics a good case-study for this analysis, as students should be able to constructively comment on their experience with feedback in such a context.

Student recordings and focus group analysis

Nine students participated in more depth, as described in the 'Methods' section. This self-selected sample is roughly representative of BASc/EBEES numbers but does not include a PEES student. With only one male participant it is very unrepresentative in gender.

#	Alias	Programme
1	Irina	BASc
2	Daniil	EBEES
3	Margarita	BASc
4	Helen	EBEES
5	Petya	EBEES
6	Diana	EBEES
7	Marta	EBEES
8	Nina	EBEES
9	Alexis	EBEES

Table 5 - Alias and programme of study of the 9 focus group participants

Data analysis

The co-investigator outside the department carried out open coding of 14 transcripts in Nvivo, with a view to allowing themes, and ultimately theory, to emerge from the data (Strauss and Corbin, 1998). This yielded the set of nodes shown in 'Table 6 - set of themes generated from transcripts'. Nvivo allowed intersecting content to be identified – for example, which mentions of 'feedback' which were also coded as 'helpful', or which mentions of 'criteria' also made reference to 'digital' aspects. Nvivo also revealed which preoccupations were concentrated in a small number of sources and which were more widespread. The nodes were then used to generate a model of what practices or aspects make assessment criteria and feedback helpful from a student point of view, and what practices or aspects don't (see 'Executive summary' and 'Key findings').

Node name	Sources	References
feedback	14	165
formative	6	14
essay	5	11
modes	6	11
summative	1	2
unhelpful	13	122
helpful	14	101
criteria or guidance	13	80
digital	13	44
Turnitin	8	29
Moodle	10	22
numeric grade	12	34
worries, trust, &c	10	33
model answer	5	19
key	8	17
Topics in Micro	3	14
project comments	5	10
other than Topics	1	8
effort	6	8
revision	4	5
debated	2	2
role	0	0
Marker or lecturer role	10	48
Student role	11	41

Table 6 - set of themes generated from transcripts

Views of assessment criteria

We need to keep in mind that, for these 2 pieces of coursework, the same set of criteria (see 'Appendix 4 – assessment criteria') were used for diverse work including essay questions, calculations, explanations, illustrations and graphs.

What makes criteria unhelpful?

There was general agreement that the criteria were too "abstract" and "generic" to be useful, which reflects the questionnaire responses of the wider cohort. This, from the focus group, is illustrative:

> Nina: me personally, I haven't looked at them before writing the coursework. I skimmed through them when [tutor name] mentioned them, but I didn't feel it useful. Because so – so abstract, for me.
>
> *[Murmurs of agreement.]*
>
> Daniil: really generic.
>
> Nina: everyone can write it on every type of coursework, and explain it in their own words.
>
> Alexis: and I think it's quite general. Like, "Show your knowledge".
>
> Margarita: *[laughs]* "Show your knowledge". Huh?
>
> Alexis: I think it's obvious that when you write something, you want to show your knowledge.
>
> Irina: yeah, and it doesn't really – you can't see how the teacher is actually going to mark your work – well, comparing to the mark-scheme – because – I know [tutor name] says that we need to explain what we're doing, that given we are writing a mathematical equation we need to explain where our variables came from. But it's never in the marking criteria.
>
> *[Murmurs of agreement.]*

Here Nina refers to the interpretation which assessment criteria demand of students and assessors in order to transfer them into judgements. However, she perceives this as slipperiness rather than nuance – "… everyone can write it on every type of coursework, and explain it in their own words".

Participants expressed doubts that the assessors referred to the criteria during marking. In Petya's opinion, "it's kind of like they read it, and they say "Meh"'". Alexis doubted assessors had time to map student work to the criteria. Daniil assumed that "after correcting essays for years, they just know what mark it is.", as did Helen: "I think they have an overall thinking about our work, and then they do the grades. I mean, that's my opinion about it. But they don't really look". When the facilitator asked if she meant that the assessors came to a numeric grade by instinct, Daniil responded "I think we all think this", which met with general agreement. From the student's comments, it seemed that this assumption of assessors' tacit knowledge reduced the validity or the benefits of assessment criteria, as a mean to communicate marking standard to students. However, as already mentioned, there is a role for criteria beyond arriving at a grade, including making a marker's tacit knowledge explicit to other

markers, improving consistency in grading through these understandings, and to communicate markers' expectations to students.

Word limit is a constraint rather than a criterion but the degree of interpretation it required from students, and the exasperation this could produce, is important to note. The word limit imposition – while generally welcomed an indicator of the level of effort and detail required - was perceived as clashing with other advice from tutors:

> Alexis: I think it's hard to find the middle way because sometimes they say there's not enough detail in your writing, but then the other hand -
>
> Daniil: - you don't have any more words.
>
> Alexis: - that you have too much words, or that you put too many irrelevant things.
>
> Diana: *[emphatically]* yeah.
>
> Alexis: so it's not easy to find the balance, I think.
>
> Daniil: all the feedback had something like, "You should have developed better this concept". But when you think of words – where?

Considering digital aspects, five students mentioned the Turnitin rubric (criteria) being difficult to find, including three who only discovered it existed after the focus group. This meant that students were potentially missing part of their feedback. Once it became apparent (from the video recordings) that students were not finding the rubric, we were able to take immediate action using the module's Moodle forum. The tutor affirmed this:

> "On Turnitin, the criteria are hidden somewhat and it is only a tick-boxing exercise: I found it more difficult to fully trust that the students would understand their scores without any personal comment added".

She mentioned that Turnitin was limited in being able to attach a comment to more than one criterion, and also raised the issue of needing to manually adjust the Moodle rubric column widths.

What makes criteria helpful?

Focus group participants had little to say about what makes criteria helpful. They seemed to feel they would understand them better if they could relate them to where they had lost marks. One proposed remedy was to give each criterion a numeric weighting and grade students for each. Marta explained:

> "I'm doing [Module Alias 03] and as my essay I got feedback just now and it's really helpful because it actually gives me a mark on my grammar, a mark on my presentation – so it's like really – like there are six or eight different points that she gave me marks – like 12 out of 15, 10 out of 15, so it's like very useful – I can see "I'm really doing bad in this, but I'm really good at this"… I could see them in advance before so I knew what I was aiming for".

Diana also raised this point:

> "… I can see that she did use the criteria to mark it, but not explicitly, you kind of had to think about it, kind of have to interpret what she said and if they would go into that criteria. I think it'd be a lot better if they explicitly used the grading rubric to tell you where you lost your points, then we can improve…".

(Both Moodle and Turnitin afford this.)

To help students pinpoint where marks were lost, another remedy proposed by two participants was for each question or sub section to have its own marking grid.

With respect to word limit, Margarita and Daniil approvingly referred to modules where they were given precise, section-by-section guidance about how to expend their word allowance. However, Margarita advocated leeway with essay questions since these were likely to vary in structure.

On the digital side, the tutor valued the ability in Moodle Assignment to give very criterion-specific feedback, which is currently impossible in Turnitin: "I thought Moodle was much better for linking comments to criteria. I really like the comment box for each criteria on the opening page". With Moodle Assignment, students easily found the rubric – in fact it was so easy to find that students unaware that there was a Turnitin Rubric referred back to the Moodle one during their summative assessment.

Assessment criteria versus model answers

Several students felt the criteria irrelevant for this diverse work and unhelpful in directing them where to focus their time and efforts. Consequently there was a long and considered discussion of the relative merits of model answers versus assessment criteria during the focus group.

Daniil advocated model answers to objective problems where your answer was either correct or incorrect: "When you have a mathematical problem you can't show, like, 'Further understanding'". Alexis suggested supplying correct graphs. Nina felt the model answers had been more valuable than the feedback on those questions: "I don't know … like after checking the solutions I can tell what I had wrong what I didn't know, but with just assessing this feedback I don't know really what should I do differently".

Essay questions were a different matter, though. Margarita pointed out, "…one person would have this structure for the essay, another one would be like… I don't know, there are so many opportunities here – it's not A Levels any more where you have to like seek to the model answer".

However, she found it worrying that,

> "…I think we were writing like a lot, a lot, a lot, and then the model answer was released and it was just a few sentences in each, and it was like "It's enough". And you're like, "I spent so much time!" – and it's not important in the coursework, because you can expand, but it's important in the exam, because if you waste too much time, on like particular, small questions, you don't have time for the big ones."

She suggested guidance about how much time and effort to make in the different parts of their work. However, Helen felt there might be something to lose here because,

> "…there's a lot of creativity in the essay. If everything is like "The introduction should be 200 words, *this* should be 200" – and then what can we really do? Like there's a part in the essay that is really nice, when you don't expect anything from this essay and you see something really surprising inside the essay, and this is why you get a first."

Petya pointed out that some highly structured types of written assignment benefited from model answers, and Daniil replied,

> "I think you can make a model answer, but it depends if you want to, as a teacher, because then all the students will probably just follow the structure and everything. Or probably just follow the thinking."

One student, in her summative feedback response, raised the possibility of model answers corresponding to each grade – this did not come up in the focus group but was resurrected by participants during their validation of this draft report.

Views of feedback

What makes feedback unhelpful?

The cryptic nature of some of the feedback was raised by Nina:

> I answered how I thought it was right and now I have just these two words: "[two words of feedback]". I don't really know how to improve it. I think that I should link those three points to the graph, but it's just me assuming, so I don't know how to improve this section, this first part.

Helen agreed: "Maybe also it could explain more what she thinks on the way we write and not only the result", as did Margarita: "...this question is just "No". Which is not particularly understandable for me.", remarking "I think those kind of comments, they reflect the thoughts of the teacher who's marking", rather than the needs of students. This included feedback phrased as questions frustrating. Nina explained:

> … when the comment is a question such as "why not use the other model?" is thought-provoking but not necessarily helps me to reach a better answer, since it might lack the reason why to use that model suggested by the question and why it would be better than the one I used.

Participants invariably found lack of direction unhelpful. Helen:

> So the only remark that I got was like "Some mistake of precision but also some good intuitions". And this is kind of like overall … like it's not really helping me to get better, I mean … yeah I can be more precise but we'll have also a word count, and yeah I have good intuitions, but why then do I have only [low mark] out of 40?

Nina's comment is also typical: "just saying that something is not logical or something is not clear or just putting just some two words feedback bubbles – I think it's not enough." Without advice, she felt very unclear about what was expected of her. Margarita agreed, "I can't say that I completely change the way I write essays after this, just because I got this feedback. It's not something super game-changing".

Some students were visibly stung, in their screencast recordings, by feedback which told them that they had not understood correctly. One voiced frustration with the lack of direction, "I'm sorry but I think once I've done it, it's pretty obvious that I thought I knew what I was doing, because I didn't do it just for nothing".

Variation in participants' responses (see the section on 'What makes feedback helpful?' for more positive views from participants) indicated that there was a qualitative variation in the feedback, and whether it was present or not. Some said they appreciated the depth and specificity while others, like Daniil, lamented the lack of it, or the lack of feedback at all:

> "… my feedback was formed exactly out of two words, both addressed for a single paragraph from my whole essay, and there was a graph which was lacking something. So basically I don't really know what I've done wrong".

The brevity was sometimes deliberate on the marker's part, since she had already expanded on the issue in the solution she provided to all students as feedback (see the 'Methods' section). However, along with the perceived patchiness, brevity could also be a symptom of marker fatigue or shortness of time, which in turn raises questions about overwork. The tutor confirmed that marking nearly 100 scripts twice in a term, each within four weeks, was very challenging indeed. Efforts were made to maintain consistency but fatigue was unavoidable. This also raises questions about student expectations, how they might interpret an absence of feedback in the context of an abundance of feedback, and how assessors can best respond, given that 'more' is not necessarily 'better', nor even practicable at current levels of resourcing. Also for consideration (but undeveloped in the focus group due to time pressures), two participants raised doubts that feedback could be useful in the context of an essay, since it's "subjective" and "it's harder for her to help us out".

Participants said they were frequently perplexed by feedback which didn't seem relevant to, or commensurate with, the mark. Helen commented, "…to some extent yeah I'm kind of disappointed because the only part that I didn't do super-well is the part that is like less explained", and later, "she said that I did like very well the part A, B, C and then like the part D it's like really really bad. And maybe it would have been great to tell me well you know like you did this great, but you could have gone in depth into this this". Diana contemplated how difficult it must be for assessors to work out what a student might have been thinking: "that's probably why they are so not very good at giving feedback about what you did – cos they don't know themselves the specifics".

The relationship between the criteria and the feedback could be onerous to figure out, as Diana illustrates:

> These comments, they don't really actually mention the part in the grading criteria that's being affected by this and how bad it is, like where it falls in that grading criteria. So I don't really

> understand where the points are lost the most. Like did I even lose a point here because the graph is too small, or is she just winding me, know what I mean?

Similarly, the relationship between the feedback and the numeric mark could be confusing:

> Daniil: … if your work is about 60, all the feedback is going to be –
>
> Diana: - really vague -
>
> Daniil: "Good work", like, -
>
> Diana: Yeah, that's true!
>
> Daniil: "Excellent analysis", "Good insight" – everything's good.
>
> Diana: and then you're like "Where are all the marks gone?!"

For one question – an essay question – no participant received more than 15 marks out of 40, which caused disappointment and confusion. Contrast between positive feedback comments and a relatively low mark frustrated some of the participants. Referring to a different module, one described their confusion when they received praise although they had achieved a lower mark than they expected. On repeating the work, they achieved a very high mark, but said they could not understand why in either case.

In general, the concept of feedback and criteria espoused by these students was largely to explain where they had lost marks and avoid this in the future. However, there were signs that this approach frustrated students when it came to the essay question. Marta was bewildered by the cohort's relative low achievement in the essay question:

> "…either we're all absolutely dumb or the question was, like, not explained well enough. But, like, we cannot be scored, like, 20 out of 25, and then get 10 out of 40 – I mean, that is just… I don't know."

Petya's account is another example:

> "I was just going to say that for Part D … I was really happy – I was like "Huh, I'm going to get a lot of points because I did exactly what she just said", but then I think I got *[under 15]* out of 20. And when I saw my feedback, I didn't see why I lost so many points, if I just – like, the structure was the same. I don't know if it was because I didn't analyse the graphs enough, or…"

Occasions like this where students thought they understood the feedback but found it did not transfer into improved marks would require further exploration. It may be that the student did successfully use the feedback, but lost marks on an aspect the feedback had not broached. However, this example does affirm the difference between the giving of feedback and its successful contribution to improvements.

Timeliness was raised. As Marta put it, "Now, after a long time – I am less motivated to actually read the comments and get into the topics with that much determination".

Considering digital aspects of unhelpful feedback, it emerged that students who had entered Turnitin via a particular route which took them to their original work but without the feedback, were prone to overlook the Grademark icon altogether and risked not finding their feedback. In Moodle, Marta and Margarita reported that the markings sometimes obscured their original work, causing them to have to open the original work outside Moodle and cross-reference with the feedback. Margarita found that downloading the file with the markings and opening it in a PDF reader resolved this (and as far as Digital Education knows, the problem has been resolved in the recent Moodle upgrade).

In Moodle the ticks and crosses and other symbols did not communicate as the tutor had hoped. The tutor had expected them to be more efficient, allowing her to give more and finer grained feedback. Marta commented, “Ticks were self-explanatory, but crosses – indicating I have done something wrong, I would like to have more explanation”. Petya mentioned that “…she would put question marks at some questions – I was like, “What does that mean?”. Moreover, because the symbols relied on positioning rather than attachment for context, it was sometimes hard to be sure precisely to which part of the work they related (Moodle Assignment has facility to improvise label lines, but this is potentially more time and effort). The tutor also found deleting markings in Moodle unintuitive, which implies a need to carefully induct markers to these assessment platforms.

What makes feedback helpful?

There was less agreement on what makes helpful feedback than what makes helpful criteria. The feedback from the Topics in Microeconomics tutor was generally warmly received. EBS participants in particular said that it contrasted favourably with other feedback given elsewhere in SSEES. Helen explained how "others, they like give a sentence about the overall thinking of the essay, or something. Whereas in Micro, they take attention with every specific part of the essay or question". Petya underlined this with reference to her summative coursework feedback,

> … for the one that counted – the, the – the other one – she gave a lot more feedback and she was explaining what was wrong and what she wrote more on. And compared to SSEES I think she does a lot more. She – like, in the other course, I don’t feel I have enough feedback – like, it doesn’t bring me anything outside my mark.

An early focus group question elicited students' the best experiences of feedback at UCL. Helen described being called to an office hour so that her tutor could find out more about how she had attained a good grade but in an unorthodox way. She valued this because it was personal and extra effort. Nina described a video in which her Statistics tutor worked through each question pointing out likely mistakes at each stage and directing students to particular points in the handouts. She valued this because it helped students understand “what we needed to do to improve, rather than just what we did wrong”, Daniil explained.

Feedback which included advice was highly valued by all participants, as much as feedback which explained a mistake. Petya praised the Topics in Microeconomics tutor's practice in a different module she had taken, where feedback had been focussed into marking grid indicating how students had done, along with summary comments:

> She gave that, which I thought was really helpful, and then a few remarks at the end, overall, and what she thought of it, and what we did wrong, what we should emphasise on, and I mentioned in my audio thing that I think all lecturers should do that, so that we can have – like, it's really easy to see what you did wrong and what you did not do so well… with the overall feedback, I think.

Daniil and Diana agreed this was best, but Margarita said that for her, ticks on a grid lacked context and specificity.

Specific, closely-contextualised inline comments which "specifically say what the mistake is, with highlights exactly those sentences or words that were wrong", gave "more specific parts of feedback on specific things that I did wrong" were also mentioned by Nina, Alexis, Petya and Helen, who said:

> I did not have any summary feedback, it was more for each part of the coursework. But it explains quite well what are the issues and what is the tutor expecting as well. Cos you also have the formative assessment corrections, so you know what the teacher was expecting us to do.

Feedback which elaborates on the issue was valued by all students. As Daniil put it, "… do not try just one or two words, just write like a sentence. Like "Here you should have looked at this", or … I don't know "This contradicts this fact" - you know". On received this kind of feedback, Diana responded appreciatively,

> … "in which case the bend is explained by a switch from a dominant substitution effect to a dominant income" - well that's good what she wrote there because that's actually … like I thought I realised it but I was still unsure and now I know so … get a better grade on a real coursework, so good to know.

Explanations of how marks had been lost, with reference to the relevant criterion in each case, were very valuable for Diana.

Considering digital aspects of helpful feedback, Daniil preferred online feedback because handwriting could be illegible. He and Marta also preferred the way Turnitin rendered comments, comparing it positively to Moodle where the comments had sometimes obscured their original work (Digital Education consider this a technical issue rather than an assessor one, and have taken steps). As mentioned, Alexis and Petya preferred the specificity and contextualisation of inline comments, as did Diana with reference to her formative feedback which came as highlighted text with MS Word comments attached, since she had submitted that format (note that Moodle now renders MS Word submissions as annotatable PDFs, if wanted). For Nina too, "the most useful is the comments that are embedded in the text. They are easy to find and gives specific feedback on the mistakes or missing things".

The tutor found bubble comments helpful. In her future use of the ticks and crosses which Moodle offers, she said she would ask students to leave a larger margin round their work.

Moodle Assignment has resolved another complication mentioned by the tutor. Since the July 2016 upgrade, Moodle Assignment (like Turnitin) brings together the different modes of feedback on a single screen.

Effort

Effort relates to assessors' and students' energies. We give effort particular attention here because it was a distinct theme in the focus groups. It probably goes without saying that assessment is a huge cognitive and emotional effort for students who are usually new to the disciplinary terrain. Fundamental to their bewilderment about the essay was their judgement about common knowledge, and this fed dilemmas about what to cite and where to expend their word limit. Diana, recording herself as she perused her formative feedback:

> "And you could provide a more formal definition of complexity. But good overall'. More formal? Well, like a text book definition? Isn't that plagiarism? Oh - I didn't even put in a definition. Well doesn't everyone know what complex is? Why do you have to define it? It's just common knowledge.

When this came up in the focus group, the facilitator pointed out that becoming good at these judgements was something students were here to learn – they were not expected to arrive with this knowledge. Unsurprisingly, participants did not seem reassured.

In times of students' perennial entreaties for more feedback, producing feedback can be a huge effort for assessors, as Petya illustrates:

> "I think the more – well it's obvious – the more the lecturer spends time – the [Module Alias 02] lecturer – [s/he] put a lot of effort and a lot of time on our coursework, and it ended up helping us a lot more."

However, she explained how this particular feedback demanded a huge effort on the part of the lecturer because "loads of people emailed [him/her] like two days before – a lot of people – and [s/he] kept replying until the morning of the exam". This level of feedback would be beyond many assessors' capabilities since they would not be allocated the resource to be this responsive.

Some participants, like Alexis and Marta, speculated that fair assessment itself might be too much effort:

> "...obviously they're rushing to read out through all the papers or through the answers I wrote – so it's like an overall impression and not like really "Oh – analysis was good but this organisation was bad" – they just give an overall "OK". Because for example the essays, like a 5 page – you don't remember the whole essay, you put down the overall impression. Or you put down the impression that you got in the last page because it was obviously the last in your memory."

The effort expected of assessors in providing feedback can contrast with the effort students expect to make in practice assessments which 'don't count':

> "I feel my mark is a bit low, cos I only get [lower mark], but it was a formative assessment so I know that for the summative one I went really in depth and so I understood my grades. But it was quite appropriate to the effort I gave to this work."

Another participant:

> "I didn't pay too much attention to this coursework as it was not assessed but I wanted to submit it and see what I can get without any preparation."

In the focus group, Helen acknowledged this attitude as widespread, but questioned it:

> …the fact that the formative was less like – students didn't pay that much attention to it, doesn't mean that's cool – because like you worked on it, and it could have helped you do more and be better for like the summative that counts on your grades and everything.

This low-effort phenomenon raises the question of whether it would be a better use of assessor energies, and more beneficial for students, if assessors were giving feedback on students' best efforts rather than supplementing low effort. However, action on this front is likely to be complicated by students' impression that the worse they do, the more attention and better feedback their work receives:

> Daniil: … If your work is really bad, then they say "OK, *this* is wrong, *this* is wrong, *this* is wrong…" But if your work is about 60, all the feedback is going to be –
>
> Diana: - really vague -

Students had the impression that summary feedback was often the same for different students, and implied that they found this impersonal. At the same time, participants, including Petya, Daniil, and Alexis, made reference to the various pressures on assessors.

Another aspect of effort is taking up feedback. Margarita illustrates the demands which processing different modes of feedback simultaneously can place on students:

> "Because kind of here it's all over the place so we have to look at questions, you have to know what you've written there, then you'd have to look at what the general answer should be … like what she had in mind when posed us those questions, it's like all over the place, so you have to look at [inaudible] and it's really difficult to comprehend all of it."

Diana also illustrated this with an exceptionally long formative screencast recording during which she conducted a complex four-way mapping process between criteria, bubble comments, summary comments, and her own work. Notwithstanding its exceptional length, at the end of the recording she remarked that she had barely skimmed the feedback and she need to do more work with it. She went on to improve her summative mark by two degree classifications, and felt that working with her formative feedback was largely responsible. Weeks later in the focus group she was the only participant who felt she could judge how well she had done in her essay before she received the mark. Diana's experience bears out research findings on the considerable demands that working with criteria and feedback places upon student, and the benefits which can accrue if students do this work.

With respect to effort with digital criteria and feedback, Daniil weighed up the relative merits of online and paper feedback, and identified the need for an internet connection as potentially limiting for markers (it may not be widely known that both Turnitin and Moodle enable marking without an internet connection).

Views of SSEES and UCL

Although our study centred on criteria and feedback, students were particularly exercised by what some viewed as the distinctly UCL practice of not using the full range of marks. Helen contended that nobody could get more than 80% for an essay, commenting,

> That's unfair because when you apply for a Masters later, you're compared to other universities, like [other University of London institution], and they do have grades, like, up, like, better than ours, and that's unfair because they don't take your candidature, like, they don't want you, when you may be better than them. Because it's strict –

One participant remarked that the marks are secretly capped. If that were widely believed it would presumably have a demotivating effect on effort.

Petya, Helen, Marta and Nina independently attributed what they viewed as superior feedback in Topics in Microeconomics to the existence of the project. Although participants were aware that the project had in fact been instigated by their assessor with a view to improving feedback, Petya commented, "it's obvious that she works on it because we're going to assess it later" and expressed a wish to broaden the benefits by running the project as a permanent fixture in which all students would participate. On the other hand, Marta critiqued the project validity, suggesting that the study be repeated with a sample of students selected after completing their module because "there's a possibility she was trying harder than she was like last year". Underlying these contributions is their concern about the persistence of current norms of feedback provision, and their hope for feedback which is more like the feedback they receive in Topics in Microeconomics.

When asked in the focus group how students might make best use of feedback, Nina immediately responded that she would initiate a dialogue with a tutor, and a discussion about bias ensued. Marta said that she had wanted to query her marks previously but "I don't even think to do this in SSEES because they know our student numbers and they know my name and I don't want to be marked down". Daniil also felt a possibility that a tutor might attempt to track him using his student number. Some participants doubted this, but others had a perception that bias was a factor in their marking.

This matter of low trust needs to be addressed, since it is a liability to the balance between UCL's chosen approach to addressing unconscious bias (anonymous marking), and the competing good of tutor and student being able to discuss the feedback, which is highly beneficial according to the research literature. It is also a liability to confidence in the assessment process, and good staff-student relations.

Discussion

Participants were alive to the dangers of over-specifying criteria leading to writing by numbers. At the same time, they often felt unsure of what was expected and the bare presence of assessment criteria was insufficient to guide them. They were therefore unsure about how to succeed. After first asking if they felt that they were getting better at making these judgements, the focus group facilitator reframed the question, “Who thinks it’s still a lottery – who has *no* idea how they’ve done on the essay question”. Only one student (Diana) didn’t raise their hand. Along with the perception of the criteria as slippery rather than helpful constraints, this points to the need for students and tutors to reach a shared understanding of the criteria for a given assessment (c.f. Bloxham et al, 2015; Reddy and Andrade, 2010).

Student doubts about tutors’ use of the assessment criteria during marking probably relate to students’ very low engagement with criteria, as expressed in the questionnaire. Students take a strong cue from their tutors, and if they do not think their assessors are making reference to the criteria we can deduce that they will see little reason to do so themselves. As such this is a concern since the criteria exist to help students comprehend the standards against which their work will be judged. Criteria are therefore important to perceptions of fairness and confidence in the marking process. This confidence is in turn important to students’ regard for tutors and the institution. It is likely to be helpful, then, to emphasise with students the role of criteria beyond guiding the marking process. If experienced expert assessors have internalised the criteria and now mark tacitly, there is still a role for the rubrics and marking grids in giving feedback - making these assessment judgements more explicit to students, and also allowing students to benchmark themselves and see where to focus their efforts.

There was an unresolved debate about an appropriate level of specificity for the criteria. Irina started a discussion about there being no reference to explanations in the assessment criteria. Nina agreed, recounting how she had included an equation without an explanation, which had prompted a helpful reminder from the assessor: "You should explain". But Petya thought this was elementary - "well we are at uni, we know how to write an essay" and called for more specificity while Diana also thought that phrases like "show your knowledge" were "a bit too vague".

In her summative feedback recording, Petya summarised why she found her feedback helpful:

> So basically what my feedback is, that it’s very different from my other feedback … we clearly have total marks and the feedback is a lot better than the other coursework. We actually know how much we have, over how much for each question, which I think is really important because we can see how well we did in one question compared to another. So yes I can see what I did wrong, which mistakes I made and what I could have avoided to get a higher grade. The correction did specify what was lacking in my answer. I felt they were a lot more … well they were specific, but they were actually noting what we did right and what we did well, and then with a little bit more information on what we could have highlighted a bit more, which I think is important because we get to select the information that we have.

At the same time, the tutor observed that the students who attended her office hour did not seem to have read the feedback she had given. Participant recordings illuminated the cognitive demands on students mapping bubble comments, criteria, summary comments and their own work. As Diana's experience in particular showed, this hard work is beneficial - but it is also daunting and time consuming, and the benefits may not be immediately obvious to students. Consequently students may not carry out this exercise. Another reason they may not carry out this exercise is if the feedback is so hard to find that they do not register its existence.

Students had no strong preference for either Moodle or Turnitin. Moodle was viewed as superior to Turnitin in terms of making the criteria discoverable and allowing criterion-specific comments (although students need instructions with both platforms). Turnitin, on the other hand, was viewed as 'better made'. Although Turnitin allows each comment to be attached to a single criterion, the tutor found this limit of one criterion restricting.

Key findings

We believe that recommendations are best generated by representatives of the group who will need to work with them, so we avoid making them here. Instead we summarise a set of principles derived from the literature and the student and tutor experiences described above.

Assessment criteria

1. Students need an opportunity to discuss, and ideally practice with, the criteria in advance, so that they and the assessors can reach a shared view of the standards by which their work will be assessed.
2. Students need to know that criteria exist and be supported to use them. Moodle Assignment is good for making rubrics salient, whereas Turnitin requires students to know to click an icon.
3. Students need support to benchmark their own work to the criteria. Moodle or Turnitin rubrics allow assessors to indicate which levels students have achieved. Moreover, Moodle allows a summary comment for each criterion.
4. Since students doubt that assessors refer to the criteria during marking, it is important to make the educational case for criteria (i.e. beyond grading) as a way of reaching a shared understanding about standards, for giving and receiving feedback, and for self/peer assessment.

Feedback

5. The feedback comments most valued by students explain the issue, make links with the assessment criteria, and include advice about what students should do next.
6. Giving feedback digitally is legible and easily accessible from any web connected device.
7. Every mode of feedback should be conspicuously communicated to students and suggestions on how to cross-reference these different modes should be provided. Some thoughts should be given to ways to facilitate access to and interpretation of all the elements of feedback provided.
8. Students need to know that digital feedback exists and how to access it. A slideshow of screenshots would allow tutors to hide and unhide slides depending on which feedback aspects they are using.

9. Participants were relatively uninterested in feedback about the parts where they had done well, instead they wanted to be told how to improve their work. Departments with good NSS scores on Assessment and Feedback recommend giving specific positive feedback as well as feedback on problematic parts (Archaeology, 2016). So it could be that positive feedback also needs to highlight suggestions for further improvements. Turnitin and Moodle Assignment both give this opportunity in their summary comment field.
10. Students need to be confident that they can approach tutors about their mark and feedback without fear of prejudice in future assessments.

Effort

11. The more feedback is dispersed between different modes, the more effortful it is for students to relate it to their own work and thinking. Where more than one mode is used, there is a need to distinguish between the purpose and content of each kind of feedback, signpost their relationships, and communicate this to students. Turnitin offers some support for cross referencing between bubble comments and criteria.
12. It would be possible to ask students to indicate on their work which mode (out of a choice of possibilities) they would like assessors to use.
13. The submission of formative assessment produced with minimal effort may impose a disproportionate burden on markers, who are likely to be commenting on mistakes that students could have corrected easily by themselves. Shorter formative assessment, group works, clearer statements of the benefits of submitting formative work may all help limiting the incidence of low-effort submissions.
14. If individual summary comments have a lot in common, consider releasing them as general feedback for the cohort, spending the saved time on more student-specific comments instead. However, this needs to be signpost clearly to facilitate students cross-referencing of sources.
15. As a group, teaching teams can organise a hands-on session with Digital Education to explore Moodle Assignment and Turnitin from the perspectives of students, markers and administrators. This exposure will help immeasurably with designing efficient, considerate processes and workflows.
16. The kind of ‘community work’ referred to by Bloxham and colleagues (2015) would be an opportunity to reach shared understandings of the roles of students and markers with respect to criteria and feedback, which would in turn help to build confidence in the assessment process.

Overall, we note that any initiative to actively engage students with marking criteria or feedback, and to provide opportunities to discuss markers' expectations and learning outcomes, is likely to be rewarding. It will help to build trust between staff and students and promote the joint engagement with marking criteria needed to reach a common understanding of marking processes and learning outcomes. To this effect, further e-learning grant could be thought, but other initiatives offered within UCL could also be exploited, UCL Changemakers in particular.

Digital assessment

Although some commented that Turnitin looked “better made”, student participants did not have strong preference for either Turnitin or Moodle one way or the other. In both they appreciated the contextualised bubble comments and comments tied to the rubric. The tutor started with more familiarity with Turnitin but developed a preference for Moodle Assignment because of its flexibility.

The findings from this project have fed into Digital Education's detailed comparison between Moodle and Turnitin at https://wiki.ucl.ac.uk/x/mgX-Ag. The following summarises the distinct benefits of each at the time of writing.

Both interfaces offer:

- submission with deadlines.
- ability to hide marking from students until a certain date.
- a feedback environment which integrates bubble comments (including reusable comments), inline annotations, general comments and rubrics / marking guides, and numeric mark.
- offline marking (in Turnitin this is restricted to an iPad app).

Distinct benefits of Moodle Assignment:

- mark anonymised work offline using Word or PDF annotation tools and bulk-upload the marks and feedback (no special hardware or software needed).
- grade and provide general feedback in a spreadsheet (downloadable from Moodle) and bulk upload back into the Assignment for automatic distribution back to students;
- general feedback and criteria comments from rubrics and marking guides are visible in the My Feedback report.
- give summary feedback for each criterion on a rubric.
- add feedback symbols, including ticks and crosses.
- add freehand drawn feedback, for example connecting lines or corrections on graphs.
- upload feedback files in any format.
- support for multiple markers in the form of a workflow and some protections to prevent marks being overwritten in error.
- support for extenuating circumstances and discretionary deadlines for individuals.
- makes the rubric (criteria or marking grid) and feedback given within it highly salient (Turnitin requires student to click an icon).

Distinct benefits of Turnitin:

- the Originality Report - checking students' sources and referencing by viewing an overlay of matched phases while marking.
- summary spoken feedback recorded directly into Turnitin (no need to save and upload).
- associate a comment with a single criterion, enabling cross referencing.

- a word count (though note this is inclusive, and may be discrepant from the word processor's word count).
- marks hidden from students by default until a set date and time (with Moodle you have to actively set this).

Noting the difficulties students had finding their feedback, and the unintuitive aspects of each technology for the marker, Digital Education has responded by improving our guidance. Departments are urged to contact their Digital Education advisor for support to familiarise colleagues with these strategically important technologies, use them to best advantage and avoid the pitfalls which can affect inexperienced users. Digital Education are also investigating ways to integrate Turnitin's Originality Report into Moodle Assignment.

References

All web resources last accessed 30 August 2016.

Black, P. & Wiliam, D., 2009. Developing the theory of formative assessment. Educational Assessment, Evaluation and Accountability, 21(1), 5-31.

Bloxham, S., Boyd, P., Orr, S., 2011. Mark my words: the role of assessment criteria in UK higher education grading practices. Studies in Higher Education 36, 655–670. doi:10.1080/03075071003777716

Bloxham, S., den-Outer, B., Hudson, J., Price, M., 2015. Let's stop the pretence of consistent marking: exploring the multiple limitations of assessment criteria. Assessment & Evaluation in Higher Education 1–16. doi:10.1080/02602938.2015.1024607

Crisp, B., 2007. Is it worth the effort? How feedback influences students' subsequent submission of assessable work. Assessment & Evaluation in Higher Education, 32(5), 571-581.

Ferrell, G., 2013. Supporting assessment and feedback practice with technology: from tinkering to transformation. Jisc.

Gardner, A. and Rees, R., 2016. Improving assessment and feedback scores. UCL Arena Exchange Seminar.

Higgins, R., Hartley, P. & Skelton, A., 2002. The Conscientious Consumer: reconsidering the role of assessment feedback in student learning. Studies in Higher Education, 27(1), 53-64.

Lipnevich, A.A. & Smith, J.K., 2009. "I really need feedback to learn:" students' perspectives on the effectiveness of the differential feedback messages. Educational Assessment, Evaluation and Accountability, 21(4), 347-367.

Newland, B., 2016. HELF UK HE Research on Electronic Management of Assessment, 2016. Available from http://www.slideshare.net/barbaranewland/helf-uk-he-research-on-electronic-management-of-assessment-2016

Orsmond, P., Merry, S. & Reiling, K., 2005. Biology students' utilization of tutors' formative feedback: a qualitative interview study. Assessment & Evaluation in Higher Education, 30 (4), 369-386.

Reddy, Y.M., Andrade, H., 2010. A review of rubric use in higher education. Assessment & Evaluation in Higher Education 35, 435–448. doi:10.1080/02602930902862859

Sadler, R., 1989. Formative assessment and the design of instructional system, Instructional Sciences, 18, 119-144.

Strauss, A. L. and Corbin J.M., 1998. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. 2nd ed. Thousand Oaks: Sage Publications.

The Economics Network, 2016. NSS Results and Comparison (2010-2016). Available from: https://economicsnetwork.ac.uk/research/nss2016

Appendices

Appendix 1 – questionnaire

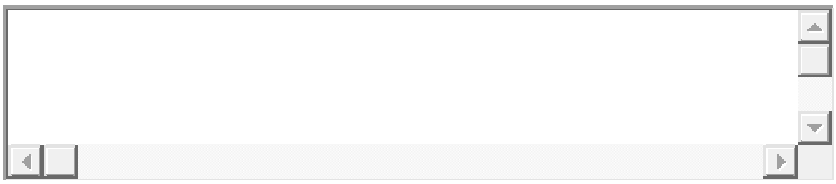
1

Response is required *

Thinking about the assessment criteria for your last piece of coursework (for any module), which of the following statements do you most agree with?

- I wasn't aware of the criteria.
- I was only vaguely aware of the criteria.
- I had difficulty understanding the criteria and difficulty relating them to my own work.
- The criteria seemed clear but in practice I had difficulty relating them to my own work.
- The criteria were clear, and I was able to relate them to my own work.

If you have further comments, type them below.



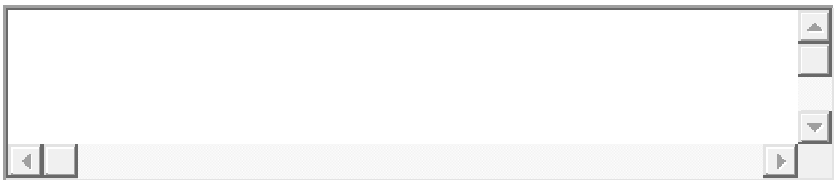
2

Response is required *

Still thinking about your last piece of coursework, which of the following **best** describes how you used the assessment criteria?

- I didn't consider the criteria.
- I considered the criteria at the start of preparing my work.
- I considered the criteria as the deadline approached.
- I considered the criteria throughout.
- I systematically tried to apply the criteria to different aspects of my work throughout.

If you have further comments, type them below.



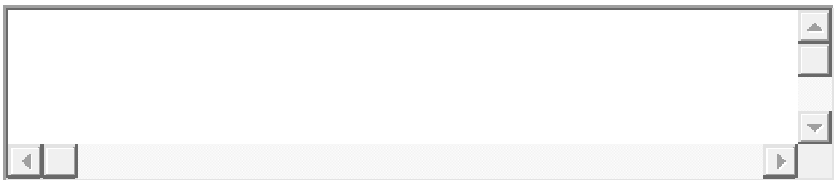
3

Response is required *

Still thinking about that last piece of coursework, how well do the following qualities describe the feedback you received?

	Strongly disagree	Disagree	Neither	Agree	Strongly agree	N/A
Informative.	○	○	○	○	○	○
Easy to understand.	○	○	○	○	○	○
Helpful for my future work.	○	○	○	○	○	○
Related to the assessment criteria.	○	○	○	○	○	○
Related to my own work.	○	○	○	○	○	○
Given at the expected time.	○	○	○	○	○	○

If you have further comments, type them below.



4

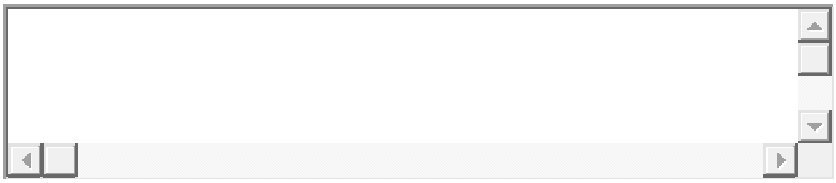
Response is required *

How did the mark you received for that coursework compare to your expectations?

- ○ Higher than I expected.
- ○ Roughly the same as I expected.

- Lower than I expected.
- I wasn't sure what to expect.

If you have further comments, type them below.



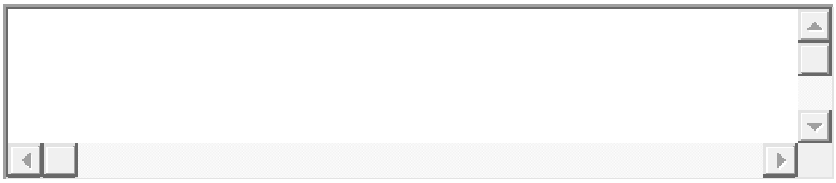
5

Response is required *

Is your gender:

- Female
- Male
- Other
- Prefer not to say

If you have any final comments, type them below.



Appendix 2 – questions for students' recordings

Questions to answer as you access your feedback

As you access your feedback, please record your answers to these questions

For **guidance** on how to make the recording, refer to the separate page linked adjacent to this one.

Just before you access the feedback:

1. Where are you and what equipment are you using? (E.g. on the bus / bedroom / computer room using a laptop / phone / tablet, UCL computer)
2. Do you have any questions, hopes, fears or other ideas in mind in advance of seeing the feedback?

As you access the feedback (you don't need to answer these in order, but try to cover them):

3. How are you getting to your feedback? Speak your route through Moodle aloud, mentioning anything you find a help or an obstacle.
4. Thinking about the comments embedded in the text - were they straightforward to find? What effect did it have on how you made sense of the feedback?
5. Thinking about the general / summary feedback - was it straightforward to find? What effect did it have on how you understood the feedback?
6. Thinking about the table containing the marking criteria - was it straightforward to find? What effect did it have on how you understood feedback?

After accessing the feedback:

7. Do you feel your mark is appropriate given the feedback you received? Why?
8. Do you think your mark is a good reflection of the quality of your work? Why?
9. Is there anything you would like to change about the feedback to make it more helpful to your future work?
10. Do you have any questions, hopes, fears or other ideas in mind after seeing this feedback?

Appendix 3 – focus group questions

1. How did the feedback for Topics (both formative and summative assessments) compare to feedback on other SSEES courses you have taken?

2. How does the feedback for Topics (both assessments) compare to the best feedback you've received so far at UCL?

3. Ideally, what do you think students can or should be doing to make best use of the marking criteria in Topics?

4. In practice, what use did you make of the criteria?

5. Ideally, what do you think students can or should be doing to make best use of the feedback you've received in Topics?

6. In practice, what use did you make of the feedback?

7. Did you feel the feedback you received was tailored to your work?

8. You used Moodle Assignment for your formative assessment and Turnitin for your summative one. How do they compare in how you have encountered the criteria and feedback?

> *Facilitator passes screenshots of Moodle Assignment and Turnitin Assignment: screencapture-moodle-assignment-student-view.png and screencapture-turnitin-rubric_student_view.png*

9. Do you have anything else to say?

Appendix 4 – assessment criteria

These were adapted from the SSEES model assessment criteria to be more particular to Topics in Microeconomics.

Criterion	**First** **(Excellent)**	**Upper Second** **(Very Good)**	**Lower Second** **(Good)**	**Third** **(Satisfactory)**	**Fail**
Knowledge Demonstrates knowledge of theoretical tools and discipline-specific vocabulary.	Accurate use of methods, tool and concepts. Clear awareness of limitations. Clear understanding of applications.	Good use of appropriate concepts and tools. Some awareness of limitations. Shows some knowledge of relevant facts	Provides a satisfactory quantity of accurate information; shows familiarity with key concepts and tools.	Shows an awareness of the relevant terminology and concepts. Provides a limited amount of information	Shows inadequate familiarity, information presented is incompetent in quality and/or accuracy.

		and statistics. Ability to draw on theoretical and applied, factual knowledge of relevance.	Some minor errors or inaccuracies.	with significant omissions. Some key concepts are missing, misunderstood or inaccurately used.	
Analysis and Interpretation Demonstrates ability to interpret the questions at hand and to analyse the information provided. Think in microeconomic terms.	Distinctive, sophisticated and focused analysis. Critical use of theory and facts and/or insightful interpretation.	Displays ability to engage critically with the question and to analyse and evaluate material effectively.	Competent understanding of the material, some critical insights presented but generally analytically "thin".	Limited understanding of the material and its significance providing a superficial or evasive answer. Wrong frame of analysis, but some relevant elements.	No evident sense of what the material means.
Structure and Argument Demonstrates ability to structure work with clarity, relevance and coherence. Ability to argue case, clear evidence of analysis and logical thoughts. Recognition of an argument limitations or alternative views. Ability to use other evidence to support	Sustained, coherent and well supported argument is presented.	A clear and coherent argument is presented. Logical answer, mostly clearly explained reasoning.	Provides an adequate, appropriate argument. Adequate level of explanations.	Poor grasp of the relevance of source material to what is being argued, weak organisation and structure. Incorrect focus. Limited level of explanations.	Incomplete answer or absence of coherent argument. Absence of explanations.

arguments and structure appropriately.					
Presentation and documentation Accurate and consistently presented footnotes and bibliographic references, accuracy of grammar and spelling, correct and clear presentation of charts/graphs/tables or other data. Appropriate and correct referencing throughout. Contextually correct handling of quotations.	Excellent presentation and full and appropriate documentation of sources (when relevant). Excellent illustrations, with clear and detailed reading keys.	High standard of presentation including appropriate documentation. Clear, relevant graphs and tables.	Satisfactory standard of presentation and grasp of appropriate principles of documentation. Clear graphs but badly explained or poor illustrations with good explanations.	Limited care and competence in documentation and presentation. Unsatisfactory or insufficient use of graphs and tables. Poor quality of illustrations.	Inadequate grasp of appropriate principles of documentation and presentation. No graphs or tables, or very poor illustrations. Explanations are missing.

Appendix 5 – quantitative analyses

Table 1: Decision to participate in Formative Assessment

	FormativeD
BASc	0.390
	0.741
Log_IntroMicro	1.887
	1.909
attend	0.662*
	0.371
consciencious	-0.053
	0.228
Constant	-12.869
	8.274

N	89
*** p<0.10, ** p<0.05,**	*** p<0.01

Table 2: Explaining students' performances

	Exam_Model4	**Exam_Model3**	**Exam_Model2**	**CW_Model2**	**Exam_Model1**	**CW_Model1**
CW	0.413**	0.490***				
	-0.171	-0.111				
FormativeD	-80.368	-86.684	9.728	196.640***	4.013*	4.526**
	-73.456	-72.392	-76.439	-68.74	-2.297	-2.162
FormCW	0.132					
	-0.222					
Log_IntroMicro	21.909	18.325	41.498***	47.264***	40.741***	21.829**
	-14.657	-13.304	-13.541	-12.177	-8.948	-8.421
FormLMicro1	17.691	21.147	-1.366	-45.916***		
	-18.279	-17.259	-18.261	-16.422		
BASc	-2.433	-2.375	-3.621	-2.541	-3.624	-2.616
	-2.989	-2.976	-3.28	-2.95	-3.26	-3.068
attend	-1.312	-1.143	-2.402	-2.568	-2.375	-1.676
	-1.662	-1.631	-1.778	-1.599	-1.732	-1.63
consciencious	-0.195	-0.217	-0.151	0.135	-0.173	-0.605
	-1.015	-1.01	-1.118	-1.005	-1.072	-1.009
Constant	-43.51	-34.519	-91.4	-116.015**	-88.368**	-14.068
	-54.275	-51.914	-55.693	-50.083	-37.953	-35.719
R-squared	0.334	0.34	0.19	0.164	0.2	0.095
N	89	89	89	89	89	89
*** p<0.10, ** p<0.05,**	*** p<0.01					

Focusing on Exam_Model1 and CW_Model1 first, we can see that the marks students have received for their final exam and summative coursework are both positively correlated with their performance in year 1 (as measured through the log of their final mark for Introduction to Microeconomics) and with their submission of a formative coursework. On average, submitting a formative coursework led to receiving 4 more marks on their exam and coursework.

Exam_Model2 and CW_Model2 replicate the analysis presented in Exam_Model1 and CW_Model1 respectively but adding an interaction term between the student's performance in the previous year and

the decision to submit a formative coursework. This shows that the students who performed less well in their first year course are significantly more likely to benefit from submitting a formative coursework.

Exam_Model3 and Exam_Model4 show that the marks obtained for the compulsory coursework strongly and positively correlate with the end-of-year exam marks, trumping all other variables in explaining the performance in the final exam. In other words, coursework marks are the only variables explaining exam performances in our cohort, past performances have become irrelevant. Further to this, the relationship between coursework and exam mark is not different for students who did or did not submit their formative coursework. In other words, whatever those who submitted their formative coursework gained from the experience is not levelled off once everyone has been through the compulsory coursework.