Automating Feedback for Computing Vivas and Presentations - A Journey

Stephen Murphy, Birmingham City University, Birmingham, UK, stephen.murphy@bcu.ac.uk

Liam Matthews-Dibbins, Birmingham City University, Birmingham, UK, liam.matthews-dibbins@bcu.ac.uk

Chris Maguire, Birmingham City University, Birmingham, UK, chris.maguire@bcu.ac.uk

Paul Shemmell, Birmingham City University, Birmingham, UK, paul.shemmell@bcu.ac.uk

ABSTRACT

Feedback and feed-forward are a critical part of the learning process. However, well-constructed, high quality, written feedback is challenging and time-consuming to produce. The growth in the number of students in higher education and ever-increasing student and management expectations has led to pressure on academics to produce higher quality feedback more rapidly. Against this backdrop, this paper introduces a system called AssessFlow that automates the process of generating accurate, individualised and repeatable feedback that is consistent across multiple markers. The paper will go on to reflect upon how the tool has been used on a number of large computing modules, its evolution and plans for future development.

CCS CONCEPTS

• Social and professional topics; • Professional topics; • Computing education; • Student assessment;

KEYWORDS

Assessment, automation, feedback, higher education, marking

ACM Reference Format:

Stephen Murphy, Liam Matthews-Dibbins, Chris Maguire, and Paul Shemmell. 2021. Automating Feedback for Computing Vivas and Presentations- A Journey. In Computing Education Practice 2021 (CEP ’21), January 07, 2021, Durham, United Kingdom. ACM, New York, NY, USA, 4 pages.

https://doi.org/10.1145/3437914.3437969

1 INTRODUCTION

Assessment feedback is a consistently contentious topic for both staff [10] and students [4]. Staff can feel frustrated by the need to provide ever more detailed feedback that is not always acted upon, while students may perceive that feedback is generic, slow to arrive, too brief or unclear [8]. Added to this, is the pressure from Higher Education (HE) management [12] to continually improve National Student Survey (NSS) scores generally and in the specific category of assessment and feedback.

Much of the assessment and feedback process remains manual, commonly an electronic analogue of the paper and pen practice that has persisted for decades. This method is coming under increased pressure when faced with less traditional assessment practices, rising student numbers and elevated expectations. This paper introduces a tool called AssessFlow designed to tackle some of these problems. In this paper, the authors set out the problems being addressed, outline the journey to develop the tool and summarise how AssessFlow works. This is followed by a discussion of future work and a short commentary on the efficacy of the tool. It is hoped that readers will identify with the issues faced and may be able to use the tools and methods described here to improve and streamline their own feedback practice.

2 BACKGROUND AND PROBLEMS

Recent changes in assessment practices have promoted the concept of authentic assessments over traditional exams and written essays [13]. These aim not only to improve realism and authenticity [7], but also academic integrity in a time when the use of contract cheating services are on the rise [3]. Whilst previous use of these assessment techniques by the authors show they are effective and broadly well-liked by students, they do generate additional workload and pressure for academic staff.

The modules for which AssessFlow was originally developed, are computing modules in a post-92 HE institution in the UK. They utilise a project-based method of assessment incorporating both theoretical and practical elements. Students build cloud-based servers over a number of incremental milestones, each of these being assessed as a viva-voce (viva) exam. The authors have observed that this “little and often” approach to a larger project promotes student engagement, aids in progress monitoring and allows regular feedback delivery. Additionally, vivas more easily facilitate a verbal feedback dialogue, critical to clarifying understanding [9, 12] as well as aiding academic integrity and inclusivity [2]. However, capturing feedback in live viva assessments can be challenging. The markers are fully engaged in the examination process and there is limited opportunity to stop and take notes.

Initially, assessment feedback on these modules was constructed manually. Feedback sheets containing the assessment criteria were completed to indicate student performance against learning outcomes and sub-criteria. Feedback was written as long-form text during the viva. This commonly varied in detail between markers and required substantial checking and reworking afterwards. Whilst video and audio recordings of these vivas proved useful for audit and reference purposes, leaving the feedback and mark allocation until the assessment had finished proved inefficient. Viva style assessments are already time-consuming, so revisiting videos later for every student was seen as impractical. As such, an approach that allows for the rapid construction of feedback during the actual assessment was sought.

In addition to the complexities of capturing assessment feedback, the consistency of feedback creates additional challenges. Both of the initial modules had a cohort of around 150 students, with multiple staff members marking the assessments. Whilst having multiple markers aids in reducing the workload and overall feedback turnaround time; the style, or even quality of the feedback generated can vary greatly between markers. In fact, the main motivation for developing this system was not quality of feedback, but the perception of inconsistency where students compared marks and feedback and encountered different styles. Standardising the quantity and quality of feedback

manually post-assessment proved time-consuming, whereas pre-assessment training and marking calibration effected only minor improvements.

While these issues are not unique to these modules, or this institution, current best practice in this area is difficult to identify. Whilst there have been reviews of automated feedback systems [11], none appear current, and many previously extant systems are no longer available. Many previously published systems also focused on easy to automate assessment styles [1] and could not be applied to the modules in question. As such, an in-house development of a more generically applicable tool was undertaken in an attempt to address these issues.

3 THE JOURNEY TO A SOLUTION

As with many home-grown tools, the initial version was a basic proof of concept. This aided in generating a feedback sheet based upon manually typed feedback, automating mark aggregation etc., but little else. This reduced the time needed for a feedback cycle, but did not solve the issue of inconsistency between markers and the risk of typographical errors in feedback.

The next innovation followed the realisation that writing feedback was largely a mechanical process. Student work is compared to the predefined marking scheme by the assessor and feedback comments are written to describe attainment, identify areas for improvement and highlight success. As such, if pre-packaged feedback statements of sufficient quality, clarity and granularity could be created in advance, markers could then rate each criterion on a scale and suitable feedback and feed-forward could be generated based upon this rating. However, in order to realise this, there first needed to be feedback statements that met these requirements.

The solution to this proved remarkably simple. One of the authors had already worked on a project where a group of computing students had been asked to review a large volume of existing assessment feedback. Following this, they were tasked with defining what they perceived as ‘ideal’ feedback characteristics. Two of the students working on this project were commissioned to take this ‘ideal’ standard and write feedback and feed-forward for a set of detailed criteria in this style. After some minor editing, this formed the basis of the feedback database for the system which will be described in the next section.

4 HOW ASSESSFLOW WORKS

This section will summarise how the system works and outline some of the features it offers. A diagrammatic representation of the system workflow can be seen in Figure 1. Each student, or group of students, will have a marking spreadsheet pre-populated with meta-data such as student ID, name and the assessment criteria mark weighting. The use of a spreadsheet means markers need no special software on their machine, reducing the need for staff training and IT support. The pre-population of key meta-data into the sheets reduces errors and decreases the time taken to mark each assessment.

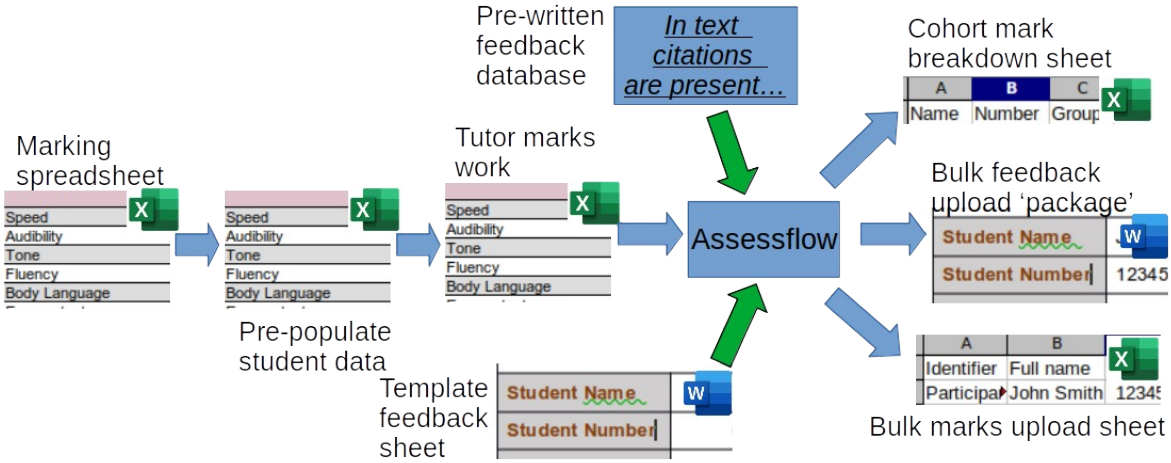

Figure 1: AssessFlow marking workflow

Figure 2 shows the basic structure of an abridged marking spreadsheet. This is created based on the published assessment marking scheme, with each criteria represented on the sheet. Each mark scheme criteria is then subdivided into micro-criteria, which are logical subdivisions of the published criteria. These were identified by speaking to markers to identify common factors they look for when assessing work and stated key sub-components in the published marking criteria. For example, a presentation could include the criteria 'oral delivery' that might include micro-criteria such as 'audibility', 'speed of delivery', 'length of delivery' etc. Each micro-criteria then has multiple indicators to represent the level of achievement in that area. The micro-criteria 'speed of delivery', may have indicators such as: 'too fast', 'too slow' and 'about right'. The appropriate option is selected by the tutor during marking. This process is then repeated for all micro-criteria for that assessment. As markers are no longer having to type feedback during a live assessment, this reduces the time needed to capture detailed feedback. Having standardised scales and micro-criteria also aids with inter-marker consistency. Numerical marks are entered for each criterion so that overall totals can be automatically computed. An additional box is available if markers want to optionally add short written comments not covered in the micro-criteria.

Team name		**Room**	
Date		**Time**	
Assessor 1		**Assessor 2**	
Student			
Student Number			
Submitted Something?		Yes	
Criteria 1	Weight (%)	7	
	Oral presentation given?	Yes	
	Microcriteria 1	Just Right	
	Notes		
	Mark (%)	64	
Criteria 2	Weight (%)	7	
	Slides used?	Yes	
	Microcriteria 1	Basic layout	
	Microcriteria 2	Just right	
	Notes	Keep formatting of slides consis	
	Mark (%)	65	

Figure 2: Abridged structure of marking spreadsheet

Once all markers have completed the marking process, all marking spreadsheets are collated into a folder. Prior to AssessFlow generating the student feedback sheets and computing overall marks, it outputs a marking spreadsheet gap analysis report. This ensures that all micro-criteria indicators and mark fields have been completed, reducing the need for manual checking. After this check and any required corrections, a more holistic second marking/verification process is carried out as per institutional requirements.

With initial checking complete, the system can be put into marking and feedback mode. AssessFlow now processes each marking spreadsheet to create individual student feedback documents in Microsoft Word format. These are based upon a standardised University feedback template, making the layout familiar to students. In operation, the system goes through each micro-criteria indicator on each sheet and matches this to the corresponding feedback and feed-forward in the database. A breakdown of the marks for the current assessment, along with a total mark are also included on the sheet. Where appropriate, the system can use information from the Virtual Learning Environment (VLE) or marking spreadsheet to cap marks for lateness or other legitimate reasons. Such capping is also reflected in the feedback and feed-forward text to advise the student of the mark reduction.

As well as showing a marks breakdown for the current assessment, student feedback suggested that a running total of their marks for a module would also be useful. Based on this, AssessFlow can optionally use a grade book export from the VLE to provide a full module mark breakdown and a running module mark total. It also allows mark inputs from other tools, one example being a peer assessment tool for group work.

To reduce the time taken for low-value tasks, the system carries out further automation of marking processes. The first such task is the naming of each student feedback sheet to allow bulk-upload to the VLE and automatic assignment to the correct student. Similarly, a CSV file containing marks for all students is output such that it can be imported into the VLE for bulk grade allocation. Both of these features greatly reduce human error and the time taken to return feedback to students. Finally, a file containing a mark breakdown for every student for the assessment (and optionally, other assessments on the whole module) is produced. This also includes submission details from the VLE (if present) and a summary of any capping applied to marks in order to allow staff to quickly look at the distribution of marks, identify areas to improve future delivery and offer information quickly and efficiently to students during meetings.

5 INITIAL FEEDBACK

This is a relatively new system and no formal evaluation of efficacy has yet been carried out. However, informal feedback from staff and students has been positive so far. As the style of feedback generated matches well with the ‘ideal’ feedback style identified by students within the School, it is likely to be as good if not better than previous practice. This is similar to the findings of previous authors on this topic [5, 6]. However, one potential area of concern is that, ironically, the feedback may be too lengthy and overly detailed.

Feedback from staff who have used the tool so far has been generally positive. One user initially expressed concern at the number of micro-criteria employed, but once they had used the system they found that it was quicker to mark than using the previous manual method. This new approach has allowed staff to deliver high-quality, consistent and detailed feedback in a timely manner. Some

staff commented that this made them feel more positive and relaxed about previously stressful assessments and that they now wish to fully adopt the system.

As intended, the system has led to a large reduction in the time needed for necessary, but low-value activities such as uploading feedback files, entering marks etc. When combined with the reduced need for manual checking of marks and feedback sheets, this time saving is significant. Whilst writing the feedback used in the database does require some thought to ensure it fits in the context of other statements it may appear with, it is no more onerous to write than standard student feedback. Given the possibilities for re-use within one, and across multiple assessments this should lead to a time saving when compared to traditional feedback methods. As the database now contains numerous entries for common criteria (referencing, presentation skills etc.), adapting it for new assessments is becoming quicker too. Additionally, feedback statements can be refined over multiple assessment cycles, allowing them to become clearer and more helpful. However, the set-up-time for small cohorts could still be prohibitive if the feedback database requires a large amount of new feedback to be written and future assessment changes preclude feedback re-use.

6 CONCLUSIONS AND NEXT STEPS

There are two key objectives to ensure the longevity of this project: evaluation and future development. Firstly, additional evaluation is planned from student and staff perspectives. Interest has been received from colleagues in other faculties and institutions, and so a cross and inter-institutional view will also be sought. It will be particularly interesting to see how non-computing staff and students perceive such a tool and the feedback it generates, and whether the use of it is practical for staff with more limited IT skills.

The system is undergoing heavy technical development with two main goals: making it easier to use for non-technical users and adding new functionality. An example of an ease-of-use development is drag-and-drop marking spreadsheet creation, currently a manual process. Feature-wise, the next version will include a report identifying the most commonly used feedback indicators. This will allow easier identification of student skill deficiencies to inform improvements to teaching practices. Looking further ahead, priorities include the ability to share the feedback database easily between users and making the system more institution-agnostic.

To conclude, AssessFlow, while designed for viva-style assessments in computing, could be utilised in any discipline for any assessment type. The larger the student cohort, the more time that can potentially be saved. Faster marking turnaround, combined with fewer errors and improvements to the consistency of feedback, may lead to an increase in student satisfaction. However, care is needed when implementing to ensure there is no disparity with standards elsewhere on a course. Its use may make other marking processes appear slow, or feedback seem terse and inconsistent. It is hoped that further development and wider use of the tool will lead to greater clarity on efficacy and additional efficiency gains.

REFERENCES

1. Michael D. Adams. 2017. Aristotle: A flexible open-source software toolkit for semi-automated marking of programming assignments. In 2017 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM), 1–6. DOI:https://doi.org/10.1109/PACRIM.2017.8121888

2. Siân Alsop and Sheena Gardner. 2019. Understanding Attainment Disparity: The Case for a Corpus-Driven Analysis of the Language used in Written Feedback Information to Students of Different Backgrounds. 7th Int. Conf. Writ. Anal. Broadening Community 3, (January 2019), 31–68.
3. Rebecca Awdry and Philip M. Newton. 2019. Staff views on commercial contract cheating in higher education: a survey study in Australia and the UK. High. Educ. 78, 4 (September 2019), 593–610. DOI:https://doi.org/10.1007/s10734-019-00360-0
4. Adrian R. Bell and Chris Brooks. 2018. What makes students satisfied? A discussion and analysis of the UK's national student survey. J. Furth. High. Educ. 42, 8 (November 2018), 1118–1142. DOI:https://doi.org/10.1080/0309877X.2017.1349886
5. Philip Denton and David McIlroy. 2018. Response of students to statement bank feedback: the impact of assessment literacy on performances in summative tasks. Assess. Eval. High. Educ. 43, 2 (February 2018), 197–206. DOI:https://doi.org/10.1080/02602938.2017.1324017
6. Philip Denton and Philip Rowe. 2015. Using statement banks to return online feedback: limitations of the transmission approach in a credit-bearing assessment. Assess. Eval. High. Educ. 40, 8 (November 2015), 1095–1103. DOI:https://doi.org/10.1080/02602938.2014.970124
7. Rachael Hains-Wesson, Vikki Pollard, Friederika Kaider, and Karen Young. 2020. STEM academic teachers' experiences of undertaking authentic assessment-led reform: a mixed method approach. Stud. High. Educ. 45, 9 (September 2020), 1797–1808. DOI:https://doi.org/10.1080/03075079.2019.1593350
8. Michael Henderson, Tracii Ryan, and Michael Phillips. 2019. The challenges of feedback in higher education. Assess. Eval. High. Educ. 44, 8 (November 2019), 1237–1252. DOI:https://doi.org/10.1080/02602938.2019.1599815
9. Katie Kerr. 2017. Exploring student perceptions of verbal feedback. Res. Pap. Educ. 32, 4 (August 2017), 444–462. DOI:https://doi.org/10.1080/02671522.2017.1319589
10. Emma Mulliner and Matthew Tucker. 2017. Feedback on feedback practice: perceptions of students and academics. Assess. Eval. High. Educ. 42, 2 (February 2017), 266–288. DOI:https://doi.org/10.1080/02602938.2015.1103365
11. Mark Shortis and Steven Burrows. 2009. A Review of the Status of Online, Semi-Automated Marking and Feedback Systems. In ATN Assessment Conference 2009: Assessment in Different Dimensions, Learning and Teaching Unit, RMIT University, RMIT University, 302–312. DOI:https://doi.org/10.13140/2.1.3576.9287
12. Jackie Tuck. 2012. Feedback-giving as social practice: teachers' perspectives on feedback as institutional requirement, work and dialogue. Teach. High. Educ. 17, 2 (April 2012), 209–221. DOI:https://doi.org/10.1080/13562517.2011.611870
13. Verónica Villarroel, Susan Bloxham, Daniela Bruna, Carola Bruna, and Constanza Herrera-Seda. 2018. Authentic assessment: creating a blueprint for course design. Assess. Eval. High. Educ. 43, 5 (July 2018), 840–854. DOI:https://doi.org/10.1080/02602938.2017.1412396